

**Learning Outcomes of the Simultaneous Exposure to Two Arabic Varieties by  
English L1 Learners of Arabic as an L2 at Different Stages of their L2 Development**

by

Mohammed Salem Alshehri

A dissertation submitted in partial fulfillment  
of the requirements for the degree of  
Doctor of Philosophy  
(Near Eastern Studies)  
in the University of Michigan  
2023

Doctoral Committee:

Professor Mohammad Alhawary, Chair  
Professor Abdulkafi Albirini, Utah State University  
Professor Marlyse Baptista  
Lecturer Said Hannouchi

Mohammed Alshehri

moshehri@umich.edu

ORCID iD: 0000-0001-8098-2769

© Mohammed Alshehri 2023

## **Dedication**

I wholeheartedly dedicate this dissertation to two remarkable people whose love for me is the purest and whom I love the most: my mother and father

## **Acknowledgements**

I attribute all praise and glory to Almighty Allāh, the Most Merciful and the Most Compassionate, for bestowing upon me and my beloved ones His countless blessings and grace, without which this significant milestone in my academic journey would not have been achieved.

In this moment of deep gratitude, I would like to acknowledge a select and distinguished group of individuals whose unwavering support, valuable guidance, and tremendous care have profoundly contributed to the accomplishment of this dissertation. First and foremost, I express my sincerest appreciation and utmost gratitude to my esteemed advisor, Professor Mohammad Alhawary, who has exemplified the true essence of a mentor, teacher, and friend throughout my doctoral pursuit. His unwavering availability to address my inquiries, even during his sabbatical, and his prompt provision of on-target comments, insightful feedback, and constructive criticism have been instrumental in refining and enhancing the quality of this dissertation. I have gained not only a wealth of knowledge from his sincere dedication and significant contributions to the field of Arabic linguistics and second language acquisition but also from his humble and respectful demeanor—which has been and will continue to be a perpetual source of inspiration across all facets of my life.

I am also deeply grateful to my esteemed committee members—Professor Marlyse Baptista, Professor Abdulkafi Albirini, and Dr. Said Hanouchi—for their support throughout the arduous journey of my dissertation. Their endless patience, constructive feedback, and steadfast belief in my abilities have left an indelible mark on my academic growth. I am truly fortunate to

have had such knowledgeable and dedicated experts on my committee, and I am immensely indebted for their contributions to my academic success.

I would also like to thank the directors and teachers of the two L2 Arabic programs from which the L2 participants were recruited. Their support and willingness to provide access to their language classes made the participant recruitment process much easier and facilitated the data collection for my study. I am also greatly thankful to all the participants who generously volunteered their time and took part in my pilot and actual studies. Their active involvement were instrumental in gathering the necessary data for my research. Without their willingness to participate, it would have been impossible to proceed with this work.

I also appreciate King Saud University for awarding me the scholarship that has allowed me to pursue my doctoral study at one of the most renowned universities across the globe. This scholarship allowed me to immerse myself in a rich academic environment, collaborate with prominent scholars, and access world-class resources. I am truly honored to have been a recipient of this scholarship. I would be remiss if I did not extend my thanks to University of Michigan's Rackham School and Department of Middle East Studies for awarding me some funding that contributed to covering some of my research-related expenses. Those aids have alleviated the financial burden and made me concentrate on my research without distractions. I am sincerely appreciative of both leading universities for their investment in my academic pursuit.

I would also like to thank all my friends and colleagues ubiquitously for their constant support and kind encouragement throughout my research journey. Special thanks go to Dr. Saeed Al Alaslal of King Abdulaziz University and Dr. Ali Al Midhwah of King Salman Global Academy for Arabic Language, for their regular check-ins on the progress of my dissertation and for their insightful suggestions of research-related papers along the way; Kamal Gasimov for his

assistance in creating the rough draft of the task pictures; James Vizthum for his help in recruiting the pilot study participants; and Ahmed Farahat for assisting me in the recruitment process of the control participants.

My immense gratitude is particularly paid to my family members. I express my sincerest appreciation, highest respect, and wholehearted love to my beloved parents for graciously allowing me to pursue my graduate study abroad and for being limitlessly patient throughout the long years I have been physically away from them. I will be forever indebted for their infinite love, unwavering support, and ceaseless prayers—without which I would not have been able to witness the successful culmination of my graduate study-abroad journey. I am also profoundly thankful to my brothers and sisters, who have always been there to provide me with every kind of support, advise, and guidance. I am really grateful for having such thoughtful siblings whose genuine care has propelled them to conceal from me some distressing news, which could have easily and majorly impacted my whole doctoral pursuit.

Last but most certainly not least, I cannot express my gratefulness and gratitude enough to my beloved wife Asma Alshehri and my heroic son Salem, who have an irreplaceable place in my heart and who are closer to it than its arteries. Asma has been genuinely my best friend. She has been a good listener, devoted supporter, and generous giver—without asking for anything in return. Her unwavering patience and love have been a source of strength and inspiration that have kept me standing and going even during the most challenging times. If she had done nothing but take care of our son, then for that only, I would remain indebted to her for eternity.

Speaking of Salem, I will make a shift in my entire discourse by acknowledging **him** directly. My beloved son, since the first moment you came into my life, you have brought immense happiness and unexplainable contentment into it. With no exaggeration, your innocent

smile, joyful laughter, and inexhaustible curiosity have been a remedy for my stress and anxiety. I will never forget the countless times when you innocently asked what can be translated as, “dad, when will you finish your work so we can play together?” This simple yet profound question will echo deeply within me forever, reminding me of the years during which you grew without me realizing it. Well, from now on, I am fully committed to being on board with you in your adventurous exploration of the long, promising, and prosperous life awaiting you, *'in šā'* *Allāh*.

## Table of Contents

<b>Dedication</b> .....	ii
<b>Acknowledgements</b> .....	iii
<b>List of Tables</b> .....	ix
<b>List of Figures</b> .....	xiv
<b>List of Appendices</b> .....	xvii
<b>List of Transcription Symbols</b> .....	xviii
<b>Abstract</b> .....	xx
<b>Chapter 1 Introduction</b> .....	1
1.1 Overview and Statement of the Problem .....	1
1.2 Importance and Focus of the Dissertation .....	4
1.3 Objectives and Questions of the Dissertation .....	7
1.4 Outline of the Dissertation .....	10
<b>Chapter 2 Literature Review</b> .....	12
2.1 Overview .....	12
2.2 The Sociolinguistic Status of Arabic .....	13
2.3 Pedagogical Approaches to Arabic as an L2 .....	16
2.3.1 Arguments for the IA .....	18
2.3.2 Counterarguments to the IA.....	24
2.4 Studies on the Learning Outcomes of the IA .....	30
2.5 Measuring L2 Development .....	36
2.5.1 Defining and measuring fluency .....	38
2.5.2 Defining and measuring accuracy .....	39
2.5.3 Defining and measuring complexity .....	40
2.5.4 Arabic studies on fluency, accuracy, and complexity.....	42
2.6 Factors Affecting L2 Development .....	47
2.6.1 Input .....	49
2.6.2 L1 transfer .....	56
2.6.3 Implicit and explicit learning .....	66
2.7 Summary and Implications of the Three SLA Factors for the Current Study .....	72
<b>Chapter 3 Methodology</b> .....	77
3.1 Overview .....	77
3.2 Participants.....	79
3.2.1 Arabic L2 program selection .....	79



3.2.2 L2 participants .....	82
3.2.3 Control participants.....	98
3.3 Target Structures .....	100
3.3.1 Tense/Aspect.....	101
3.3.2 Negation.....	108
3.3.3 Subjunctive subordination .....	111
3.3.4 Cause–effect subordination.....	114
3.4 Materials .....	117
3.4.1 Instructional input of the target structures .....	117
3.4.2 Tasks .....	130
3.5 Data Collection .....	132
3.6 Data Coding .....	136
3.6.1 General coding criteria.....	136
3.6.2 Detailed coding criteria.....	138
3.7 Data Analysis .....	146
3.7.1 Fluency.....	146
3.7.2 Accuracy .....	146
3.7.3 Complexity.....	147
3.7.4 Overall Proficiency .....	147
3.8 Summary .....	148
<b>Chapter 4 Results.....</b>	<b>150</b>
4.1 Overview .....	150
4.2 Fluency.....	152
4.3 Accuracy.....	158
4.3.1 Tense .....	158
4.3.2 Negation.....	191
4.4 Complexity.....	214
4.4.1 Subjunctive subordination .....	215
4.4.2 Cause–effect subordination.....	230
4.5 Overall Proficiency .....	243
4.6 Summary .....	246
<b>Chapter 5 Discussion and Conclusion.....</b>	<b>250</b>
5.1 Overview.....	250
5.2 Discussion of the Research Questions .....	250
5.2.1 RQ1: Nature of the language output of the IPG and UIPG .....	251
5.2.2 RQ2: Factors affecting the performance of the IPG and UIPG .....	267
5.3 Conclusion and Pedagogical Implications .....	293
5.4 Limitations .....	297
<b>Appendices.....</b>	<b>299</b>
<b>Bibliography .....</b>	<b>338</b>

## List of Tables

<b>Table 3.1</b> The Edition and Parts of <i>Al-Kitāb</i> Used in the IP and UIP, as well as the Number of Units Covered at Each Proficiency Level.....	82
<b>Table 3.2</b> General Demographic Information of the IPG and UIPG at Each Proficiency Level .	88
<b>Table 3.3</b> Time Duration of Exposure to CA Input among the IPG and UIPG Across All Proficiency Levels .....	92
<b>Table 3.4</b> Demographic Information of the Control Participants.....	100
<b>Table 3.5</b> Timing and Frequency of the Past Tense in the Instructional Textbooks used by the IPG and UIPG.....	119
<b>Table 3.6</b> Timing and Frequency of the Present Tense in the Instructional Textbooks Used by the IPG and UIPG.....	121
<b>Table 3.7</b> Timing and Frequency of the Past Negation in the Instructional Textbooks Used by the IPG and UIPG.....	123
<b>Table 3.8</b> Timing and Frequency of the Verbal Present Negation in the Instructional Textbooks Used by the IPG and UIPG.....	125
<b>Table 3.9</b> Timing and Frequency of the Subjunctive Subordination in the Instructional Textbooks Used by the IPG and UIPG.....	127
<b>Table 3.10</b> Timing and Frequency of the Cause–Effect Subordination with the Conjunction <i>li?anna</i> “because” and the EA Conjunction <i>ʕašān/ʕalašān</i> “because” in the Instructional Textbooks Used by the IPG and UIPG.....	129
<b>Table 4.1</b> Descriptive Statistics for the Fluency of the IPG, UIPG, and CG in the Past and Present Tense Narrative Tasks.....	153
<b>Table 4.2</b> Results of t-tests on the Fluency W/M Means of the CG and Each Proficiency Level Within Each L2 Group in the Past Tense Narrative Task.....	157
<b>Table 4.3</b> Results of t-tests on the Fluency W/M Means of the CG and Each Proficiency Level Within Each L2 Group in the Present Tense Narrative Task.....	157
<b>Table 4.4</b> Descriptive Statistics for the Accuracy of the IPG, UIPG, and CG on the Past Tense in the 3.s.m and 3.s.f Contexts .....	159
<b>Table 4.5</b> Error Patterns Produced by the Participants in the 3.s.m Past Tense Context (Error Type Tokens/Total Errors).....	161
<b>Table 4.6</b> Error Patterns Produced by the Participants in the 3.s.f Past Tense Context (Error Type Tokens/Total Errors).....	164

<b>Table 4.7</b> Results of t-Tests on the Accuracy Mean Scores of the CG and Each Proficiency Level Within Each L2 Group on the Past Tense in the 3.s.m Context .....	172
<b>Table 4.8</b> Results of t-Tests on the Accuracy Mean Scores of the CG and Each Proficiency Level Within Each L2 Group on the Past Tense in the 3.s.f Context .....	173
<b>Table 4.9</b> Descriptive Statistics for the Accuracy of the IPG, UIPG, and CG on the Present Tense in the 3.s.m and 3.s.f Contexts .....	175
<b>Table 4.10</b> Error Patterns Produced by the Participants in the 3.s.m Present Tense Context (Error Type Tokens/Total Errors).....	178
<b>Table 4.11</b> Error Patterns Produced by the Participants in the 3.s.f Present Tense Context (Error Type Tokens/Total Errors).....	182
<b>Table 4.12</b> Results of t-Tests on the Accuracy Mean Scores of the CG and Each Proficiency Level Within Each L2 Group on the Present Tense with the 3.s.m agreement form .....	189
<b>Table 4.13</b> Results of t-Tests on the Accuracy Mean Scores of the CG and Each Proficiency Level Within Each L2 Group on the Present Tense with the 3.s.f agreement form .....	189
<b>Table 4.14</b> Descriptive Statistics for the Accuracy of the IPG, UIPG, and CG on the Past Negation .....	191
<b>Table 4.15</b> Error Patterns Produced by the Participants in the Past Negation (Error Type Tokens/Total Errors).....	194
<b>Table 4.16</b> Results of t-Tests on the Accuracy Mean Scores of the CG and Each Proficiency Level Within Each L2 Group on the past negation.....	201
<b>Table 4.17</b> Descriptive Statistics for the Accuracy of the IPG, UIPG, and CG on the Verbal Present Negation .....	202
<b>Table 4.18</b> Error Patterns Produced by the Participants in the Verbal Present Negation (Error Type Tokens/Total Errors).....	205
<b>Table 4.19</b> Results of t-Tests on the Accuracy Mean Scores of the CG and Each Proficiency Level Within Each L2 Group on the Verbal Present negation .....	213
<b>Table 4.20</b> Descriptive Statistics for the Subjunctive Subordination Data of the Participants in the IPG, UIPG, and CG.....	216
<b>Table 4.21</b> Production of the Simple Verbal Noun Structure Instead of the Complex Subjunctive Subordination Structure by the IPG, UIPG, and CG .....	219
<b>Table 4.22</b> Error Patterns Produced by the Participants in the Subjunctive Subordination (Error Type Tokens/Total Errors).....	220
<b>Table 4.23</b> Results of t-Tests on the Means of the CG and Each Proficiency Level Within Each L2 Group on the Error-Free Production of Subjunctive Subordination .....	228
<b>Table 4.24</b> Descriptive Statistics for the Cause-Effect Subordination Data of the IPG, UIPG, and CG .....	230
<b>Table 4.25</b> Production of the Simple <i>ʔidāfah</i> construction Instead of the Cause–Effect Subordination Structure by the IPG, UIPG, and CG .....	233

<b>Table 4.26</b> Error Patterns Produced by the Participants in the Cause–Effect Subordination (Error Type Tokens/Total Errors).....	234
<b>Table 4.27</b> Results of t-Tests on the Mean Scores of the CG and Each Proficiency Level Within Each L2 Group on the Error-Free Production of the Cause–Effect Subordination .....	242
<b>Table 4.28</b> Overall Proficiency Means of the IPG, UIPG, and CG .....	244
<b>Table 4.29</b> Result of the t-Tests on the Overall Proficiency Mean Scores of the CG and Each Proficiency Level Within Each L2 Group .....	246
<b>Table 1f</b> Tukey’s Comparisons for the Fluency Mean Scores in the Past Narrative Task Based on the L2 Proficiency Levels (First Year, Second Year, and Third year) .....	327
<b>Table 2f</b> Pairwise Comparisons for the Fluency Mean Scores of the Three Proficiency Levels (First year, Second year, and Third year) Within Each L2 Group (IPG and UIPG) in the Past Narrative Task.....	327
<b>Table 3f</b> Tukey’s Comparisons for the Fluency Mean Scores in the Present Narrative Task Based on the L2 Proficiency Levels (First Year, Second Year, and Third year) .....	327
<b>Table 4f</b> Pairwise Comparisons for the Fluency Mean Scores of the Three Proficiency Levels (First Year, Second Year, and Third Year) Within Each L2 Group (IPG and UIPG) in the Present Narrative Task.....	328
<b>Table 1g</b> Tukey’s Comparisons for the Accuracy on the Past Tense Inflected for the 3.s.m Form Based on the L2 Proficiency Levels (First Year, Second Year, and Third Year).....	329
<b>Table 2g</b> Pairwise Comparisons for the Accuracy of the L2 Groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Past Tense Inflected for the 3.s.m Form .....	329
<b>Table 3g</b> Pairwise Comparisons for the Accuracy of the Three Proficiency Levels (First Year, Second Year, and Third Year) Within Each L2 Group (IPG and UIPG) on the Past Tense Inflected for the 3.s.m Form .....	329
<b>Table 4g</b> Tukey’s Comparisons for the Accuracy on the Past Tense Inflected for the 3.s.f Form Based on the L2 Proficiency Levels (First Year, Second Year, and Third Year).....	330
<b>Table 5g</b> Pairwise Comparisons for the Accuracy of the L2 Groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Past Tense Inflected for the 3.s.f Form .....	330
<b>Table 6g</b> Pairwise Comparisons for the Accuracy of the Three Proficiency Levels (First Year, Second Year, and Third Year) Within Each L2 Group (IPG and UIPG) on the Past Tense Inflected for the 3.s.f Form .....	330
<b>Table 1h</b> Tukey’s Comparisons for the Accuracy on the Present Tense Inflected for the 3.s.m Form Based on the L2 Proficiency Levels (First year, Second Year, and Third Year).....	331
<b>Table 2h</b> Pairwise Comparisons for the Accuracy of the L2 Groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Present Tense Inflected for the 3.s.m Form .....	331

<b>Table 3h</b> Pairwise Comparisons for the Accuracy of the Three Proficiency Levels (First year, Second year, and Third Year) Within Each L2 Group (IPG and UIPG) on the Present Tense Inflected for the 3.s.m Form .....	331
<b>Table 4h</b> Tukey’s Comparisons for the Accuracy on the Present Tense Inflected for the 3.s.f Form Based on the L2 Proficiency Levels (First year, Second Year, and Third Year) .....	332
<b>Table 5h</b> Pairwise Comparisons for the Accuracy of the L2 Groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Present Tense Inflected for the 3.s.f Form .....	332
<b>Table 6h</b> Pairwise Comparisons for the Accuracy of the Three Proficiency Levels (First year, Second year, and Third Year) Within Each L2 Group (IPG and UIPG) on the Present Tense Inflected for the 3.s.f Form .....	332
<b>Table 1i</b> Pairwise Comparisons for the Accuracy of the L2 groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Past Negation .....	333
<b>Table 2i</b> Tukey’s Comparisons for the Accuracy on the Verbal Present Negation Based on the L2 Proficiency Levels (First Year, Second Year, and Third Year) .....	333
<b>Table 3i</b> Pairwise Comparisons for the Accuracy of the L2 groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Verbal Present Negation .....	333
<b>Table 4i</b> Pairwise Comparisons for Accuracy of the Proficiency Levels (First Year, Second Year, and Third Year) Within Each L2 Group (IPG and UIPG) on the Verbal Present Negation .....	334
<b>Table 1j</b> Tukey’s Comparisons for the Mean Scores of the L2 participants on the Subjunctive Subordination Based on the L2 Proficiency Levels (First Year, Second Year, and Third Year) .....	335
<b>Table 2j</b> Pairwise Comparisons for the Mean Scores of the L2 Groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Subjunctive Subordination .....	335
<b>Table 3j</b> Pairwise Comparisons for the Mean Scores of the Proficiency levels (First Year, Second Year, and Third Year) Within Each L2 Group (IPG and UIPG) on the Subjunctive Subordination .....	335
<b>Table 4j</b> Tukey’s Comparisons for the Mean Scores of the L2 Participants on the Cause–Effect Subordination Based on the L2 Proficiency Levels (First Year, Second Year, and Third Year) .....	336
<b>Table 5j</b> Pairwise Comparisons for the Mean Scores of the L2 Groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Cause–Effect Subordination .....	336
<b>Table 6j</b> Pairwise Comparisons for the Mean Scores of the Proficiency Levels (First Year, Second Year, and Third Year) Within Each L2 Group (IPG and UIPG) on the Cause–Effect Subordination .....	336
<b>Table 1k</b> Tukey’s Comparisons for the Overall Proficiency Mean Scores of the L2 Participants Based on the L2 Proficiency Levels (First Year, Second Year, and Third Year) .....	337
<b>Table 2k</b> Pairwise Comparisons for the Overall Proficiency Mean Scores of the L2 Groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) .....	337

**Table 3k** Pairwise Comparisons for the Overall Proficiency Mean Scores of the Three Proficiency Levels (First Year, Second Year, and Third Year) Within Each L2 Group (IPG and UIPG)..... 337

## List of Figures

<b>Figure 3.1</b> Number Distribution of the IPG and UIPG at Each Proficiency Level.....	88
<b>Figure 3.2</b> Gender Distribution of the IPG and UIPG at Each Proficiency Level.....	89
<b>Figure 3.3</b> Age Means of the IPG and UIPG at Each Proficiency Level.....	90
<b>Figure 3.4</b> Code-Switching Frequency among First-Year Participants in the IPG and UIPG.....	93
<b>Figure 3.5</b> Code-Switching Frequency among Second-Year Participants in the IPG and UIPG	93
<b>Figure 3.6</b> Code-Switching Frequency among Third-Year Participants in the IPG and UIPG ...	94
<b>Figure 3.7</b> Proficiency Self-Rating Responses by First-Year Participants in the IPG and UIPG	95
<b>Figure 3.8</b> Proficiency Self-Rating Responses by Second-Year Participants in the IPG and UIPG .....	96
<b>Figure 3.9</b> Proficiency Self-Rating Responses by Third-Year Participants in the IPG and UIPG .....	96
<b>Figure 4.1</b> Fluency (W/M) Means of the IPG, UIPG, and CG in the Past Tense Narrative Task .....	154
<b>Figure 4.2</b> Fluency (W/M) Means of the IPG, UIPG, and CG in the Present Tense Narrative Task.....	154
<b>Figure 4.3</b> Boxplot for the Accuracy of the IPG, UIPG, and CG on the Past Tense with the 3.s.m Agreement Form .....	168
<b>Figure 4.4</b> Boxplot for the Accuracy of the IPG, UIPG, and CG on the Past Tense with the 3.s.f Agreement Form .....	168
<b>Figure 4.5</b> Boxplot for the Accuracy of the IPG, UIPG, and CG on the Present Tense with the 3.s.m Agreement Form .....	186
<b>Figure 4.6</b> Boxplot for the Accuracy of the IPG, UIPG, and CG on the Present Tense with the 3.s.f Agreement Form .....	186
<b>Figure 4.7</b> Boxplot for the Accuracy of the IPG, UIPG, and CG on the Past Negation .....	200
<b>Figure 4.8</b> Boxplot for the Accuracy of the IPG, UIPG, and CG on the Verbal Present Negation .....	212
<b>Figure 4.9</b> Means of the IPG, UIPG, and CG on the Error-Free Production of the Subjunctive Subordination .....	227
<b>Figure 4.10</b> Means of the IPG, UIPG, and CG on the Error-Free Production of the Cause–effect Subordination .....	240

<b>Figure 4.11</b> Boxplot for the Overall Proficiency Means of the IPG, UIPG, and CG .....	244
<b>Figure 1c</b> The Past Tense and Its Negation in the 3.s.m Context: The Vacation Went from the 5th to the 17th of Last Month .....	304
<b>Figure 2c</b> The Past Tense and Its Negation in the 3.s.f Context: The Vacation Went from the 10th to the 22nd of Last Month.....	305
<b>Figure 1d</b> The Present Tense and Its Negation in the 3.s.m Context: The Daily Routine of a Male Character .....	306
<b>Figure 2d</b> The Present Tense and Its Negation in the 3.s.f Context: The Daily Routine of a Female Character .....	307
<b>Figure 1e</b> Subjunctive Subordination: What Can the Girl Do/What Is She Able to Do in Each Picture? .....	308
<b>Figure 2e</b> Cause–Effect Subordination: Describe What the Man Did in the First Picture and Explain Why He Did That Based on the Second Picture.....	309
<b>Figure 3e</b> Manipulation Task Distractor 1: What Is the Girl Doing in Each Picture? .....	310
<b>Figure 4e</b> Cause–Effect Subordination: Describe the Situation of the Man in the First Picture and Explain Why He Is Like That Based on the Second Picture .....	311
<b>Figure 5e</b> Subjunctive Subordination: What Is He Thinking to Do in Each Picture? .....	312
<b>Figure 6e</b> Manipulation Task Distractor 2: Describe What They Are Doing in Each Picture ..	313
<b>Figure 7e</b> Cause–Effect Subordination: Describe the Situation of the Little Girl in the First Picture and Explain Why She is Like That or Why She Feels Like That Based on the Second Picture .....	314
<b>Figure 8e</b> Subjunctive Subordination: What Does the Woman Refuse to Do in Each Picture? ..	315
<b>Figure 9e</b> Manipulation Task Distractor 3: What Is the Man Doing in Each Picture? .....	316
<b>Figure 10e</b> Cause–Effect Subordination: Describe the Situation of the Woman in the First Picture and Explain Why She Is Like That or Why She Feels Like That Based on the Second Picture .....	317
<b>Figure 11e</b> Subjunctive Subordination: What Does the Man Want to Do in Each Picture?.....	318
<b>Figure 12e</b> Cause–Effect Subordination: Describe the Situation of the Woman in the First Picture and Explain Why She Is Like That or Why She Is in That Situation Based on the Second Picture .....	319
<b>Figure 13e</b> Subjunctive Subordination: What Does the Woman Wish to Do in Each Picture? ..	320
<b>Figure 14e</b> Manipulation Task Distractor 4: Describe What They Are Doing in Each Picture and Where These Two Actions Are Taking Place.....	321
<b>Figure 15e</b> Cause–Effect subordination: Describe What the Little Boy Is Doing and Explain Why He Is Doing That Based on the Second Picture .....	322
<b>Figure 16e</b> Subjunctive Subordination: What Is He Trying to Do in Each picture?.....	323



**Figure 17e** Cause–Effect Subordination: Describe Where the Woman Went in the First Picture and Explain Why She Went There Based on the Second Picture. .... 324

**Figure 18 e** Manipulation Task Distractor 5: Describe What She Is Doing in Each picture ..... 325

**Figure 19e** Cause–Effect Subordination: Describe What Happened in the First Picture and Explain Why It Happened Based on the Second Picture ..... 326

## List of Appendices

<b>Appendix A</b> L2 Participant Language History Questionnaire .....	299
<b>Appendix B</b> Control Participant Language History Questionnaire .....	302
<b>Appendix C</b> The Narrative Task of the Past Tense and Its Negation .....	304
<b>Appendix D</b> The Narrative Task of the Present Tense and Its Negation .....	306
<b>Appendix E</b> The Manipulation Task of the Subjunctive Subordination and Cause–Effect Subordination .....	308
<b>Appendix F</b> Tables for Follow-up Comparisons on the Fluency Means of the L2 Participants	327
<b>Appendix G</b> Tables for Follow-up Comparisons on the Accuracy of the L2 Participants on the Past Tense .....	329
<b>Appendix H</b> Tables for Follow-up Comparisons on the Accuracy of the L2 Participants on the Present Tense .....	331
<b>Appendix I</b> Tables for Follow-up Comparisons on the Accuracy of the L2 Participants on the Past and Present Negation.....	333
<b>Appendix J</b> Tables for Follow-up Comparisons on the Means of the L2 Participants on the Subjunctive and Cause-effect Subordinations .....	335
<b>Appendix K</b> Tables for Follow-up Comparisons on the Overall Proficiency Means of the L2 Participants.....	337

## List of Transcription Symbols

All Arabic examples presented in this dissertation are transcribed based on the International Phonetic Alphabet (IPA). A list of all sound symbols in Modern Standard Arabic (MSA) is provided below, following the transcription conventions outlined in previous works by Alhawary (2009a, 2011, 2019). It should be noted that titles of Arabic books and names of Arab authors are transcribed based on the conventional transliteration system commonly adopted by Arabists. This system employs symbols such as /ʾ/, /ʿ/, /th/, /dh/, /kh/, and /gh/, which correspond to the IPA symbols /ɾ/, /ʕ/, /θ/, /ð/, /x/, and /ç/, respectively. Additionally, three sounds exclusively produced in Egyptian Arabic (EA), the variety under investigation, were not included in the list. Those are the voiced velar stop /g/, the diphthong /ō/, and the diphthong /ē/, which are realized in MSA as the voiceless uvular stop /q/, the diphthong /aw/, and the diphthong /ay/, respectively. Furthermore, it is worth mentioning that some participants produced certain MSA sounds in EA-like data samples, including /j/, /θ/, /ð/, /z/, /q/ which are generally produced in EA as /g/, /t/, /z/, /z/, and /ʔ/, respectively. Conversely, in some MSA-like data samples, the MSA /q/ and /j/ sounds were sometimes produced as the EA /ʔ/ and /g/ sounds, respectively (e.g., *ʔahwa* “coffee” and *giddan* “very” instead of *qahwa* and *jiddan*). Moreover, the definite article *ʔal-* “the” is fully transcribed in all provided data samples.

Arabic Sound	IPA	Sound Description
ب	b	voiced bilabial stop
ت	t	voiceless alveolar stop
ث	θ	voiceless interdental fricative
ج	j	voiced palato-alveolar affricate

ح	ħ	voiceless pharyngeal (epiglottal) fricative
خ	x	voiceless velar/uvular fricative
د	d	voiced alveolar stop
ذ	ð	voiced interdental fricative
ر	r	voiced alveolar tap/trill
ز	z	voiced alveolar fricative
س	s	voiceless alveolar fricative
ش	ʃ	voiceless palato-alveolar fricative
ص	ʂ	voiceless alveolar fricative emphatic
ض	ɖ	voiced alveolar stop emphatic
ط	ʈ	voiceless alveolar stop emphatic
ظ	ʐ	voiced interdental fricative emphatic
ع	ʕ	voiced pharyngeal (epiglottal) fricative
غ	ɣ	voiced velar/uvular fricative
ف	f	voiceless labiodental fricative
ق	q	voiceless uvular stop
ك	k	voiceless velar stop
ل	l	voiced alveolar lateral
م	m	voiced bilabial nasal
ن	n	voiced alveolar nasal
هـ	h	voiceless glottal fricative
و	w	voiced bilabial velar glide
ي	y	voiced palatal glide
ء	ʔ	voiceless glottal stop
يّي	yy	geminate of y
وؤ	ww	geminate of w
اَ	a	short front back low
اِ	ā	long front back low
اُ	u	short high back rounded
و	ū	long high back rounded
ي	i	short high front unrounded
ي	ī	long high front unrounded
اِي	ay	diphthong
اُو	aw	diphthong

## **Abstract**

The efficacy of the “integrated approach,” which calls for the simultaneous integration of two varieties in teaching Arabic as a second language (L2), has not yet been rigorously examined, and the very few available studies suffer from many methodological limitations. The current study sought to bridge this gap in the literature by comparing the language output (in terms of fluency, accuracy, complexity, and overall proficiency) of 57 English first language (L1) learners of Arabic as an L2, belonging comparably to two groups: the first consisted of those instructed in Modern Standard Arabic (MSA) and Egyptian Arabic (EA) simultaneously, the integrated program group (IPG), and the second consisted of those instructed in MSA only, the unintegrated program group (UIPG). The participants of both groups were further divided cross-sectionally based on their three proficiency levels: first year, second year, and third year. Their performance was additionally compared to that of a control group (CG) consisting of nine native speakers of EA.

Spontaneous obligatory-in-context data were elicited from the participants who performed three oral picture-based description tasks: two narrative tasks and one manipulation task. Fluency was measured by counting words produced per minute (W/M). Accuracy was examined in their production of the past tense, present tense, past negation, and verbal present negation. Complexity was measured by their error-free production of the subjunctive and cause–effect subordinations. The overall proficiency was gauged by adding up the scores attained in the above-mentioned three constructs.

The findings revealed that both L2 groups produced comparable W/M when narrating past and present events. However, in other areas examined, the UIPG maintained a significant advantage over the IPG: (a) at the first-year and second-year levels on the past tense, (b) at the third-year level on the present tense, (c) at the first-year and third-year level on the past negation, (d) at the second-year level and (near-significantly) at the third-year level on the verbal present negation, (e) at the second-year and third-year levels on the subjunctive and cause–effect subordinations, and (f) at the second-year and third-year levels on the overall proficiency. The advantage of the UIPG over the IPG was further supported by the within-group analyses. The UIPG improved significantly in the long run—particularly at the third-year level—in their fluency, in all target structures (except for the past tense with the third-person singular feminine form and past negation), and in their overall proficiency. In contrast, the IPG did not show any significant improvement—particularly from the first-year to third-year levels—in their fluency, in all target structures in which MSA and EA diverge, and in their overall proficiency. Additionally, only the third-year UIPG participants were statistically on par with the CG in their accuracy on the past tense and present tense (in the third-person singular masculine context), as well as the verbal present negation.

The strong implication of these findings is that the simultaneous exposure to two varieties from the outset, without sufficient recycling, visual presentation, and explicit instruction of their structures in the input, is not beneficial to Arabic L2 learners and might be detrimental to their developmental acquisition paths. Such an instructional practice can cause linguistic dissonance, hinder comprehension, and obstruct facilitative L1 transfer. Thus, it is pedagogically recommended to initially teach one variety for the first 2 years of the learning process.

## **Chapter 1 Introduction**

### **1.1 Overview and Statement of the Problem**

Arabic as a second language (L2) has been taught and learned in the United States even before signing the Declaration of Independence. In the second half of the 17th century, Harvard and Yale became the first two American universities to incorporate Arabic into their curricula, primarily for theological, philological, and archaeological purposes. However, Arabic as an area of study did not receive significant attention from researchers, agencies, or authorities in the United States until the post-World War II era, when learning Arabic became one of the most studied languages nationwide, driven by academic, political, and economic factors (McCarus, 1987, 1992; Ryding, 2006, 2018). This is supported by the Modern Language Association (MLA), which has been conducting systematic surveys on undergraduate and graduate enrollments in languages other than English across all 50 states of America. According to the most recent survey in 2016, Arabic ranked the eighth most studied language, with 31,554 enrollments, experiencing a 23.9% growth between 2006 and 2016 (Looney & Lusin, 2019).

The substantially growing interest in learning Arabic has propelled Arabic L2 scholars and pedagogues to discuss how to teach the language more efficiently and effectively—and hence leading L2 learners (L2ers) to acquire it properly. However, a fundamental and pressing issue, potentially perceived as an obstacle, pertains to the diglossic nature of Arabic—or the existence of a continuum of Arabic varieties—and thus how to introduce them to Arabic L2ers (e.g., Al-Batal, 1992, 1995, 2018; Alhawary, 2013, 2018, 2021; Ryding, 1991, 2006, 2018;

Wahba, 2023; Younes, 1990, 2015). To illustrate, in all Arabic-speaking countries, at least two varieties coexist: *al-fuṣḥā* “Modern Standard Arabic” (MSA) and *al-‘āmiyyah* “Colloquial Arabic” (CA). The former is considered the official language of Arabic-speaking countries and the variety used in literacy and formal venues, whereas the latter is considered the informal variety employed in everyday spoken communication.

Traditionally, most Arabic L2 programs in the United States and elsewhere have focused on teaching MSA *only* due to its status as the official variety across the Arab world and its association with literacy. The MSA-only approach was nearly the only instructional practice in all Arabic programs until the post-World War II era, when some pedagogical approaches to second language acquisition (SLA) gained popularity. Those include the audiolingual approach, which aims to equip L2 learners with proficiency in the listening and speaking skills—in addition to the reading and writing skills (Ryding, 2018)—and the communicative approach, which endorses not only linguistic competence but also sociolinguistic competence and authentic language use (Al-Batal, 1992, 2018; Younes, 2015).

Due to the instructional challenge presented with the existence of multiple Arabic varieties, Arabic L2 programs and pedagogues have adopted at least two different instructional approaches, each of which aims at developing proficiency in MSA and at least one CA (e.g., Alhawary, 2021; Eisele, 2018). The first common approach calls for introducing MSA first then CA—typically Egyptian Arabic (EA) or Levantine Arabic (LA). However, this approach has faced criticism from those who argue that it fails to reflect the sociolinguistic reality of the Arab world, where native speakers regularly code-switch between MSA and CA on a daily basis (e.g., Al-Batal, 1992, 1995, 2018; Al-Batal & Glakes, 2018; Belanp, 2006; Hashem-Aramouni, 2011; Huntley, 2018; Husseinali, 2006; Isleem, 2018; Leddy-Cecere, 2018; Nassif, 2018; Palmer,



2007, 2008; Ryding, 1991; Shiri, 2013; Shiri & Joukhadar, 2018; Younes, 1990, 2015). To this effect, an alternative approach, called the “integrated approach” (IA), has emerged. It calls for the importance of teaching MSA and CA *side by side* from the outset of the teaching and learning processes of Arabic as an L2 (Al-Batal, 1992, 1995, 2018; Younes, 1990, 2015).

The rationale behind the IA is threefold. The first rationale is that it reflects the diglossic reality of the Arab-speaking world and creates ample opportunities for Arabic L2ers to receive authentic input (e.g., Al-Batal, 1992, 1995, 2018; Shiri, 2013, 2018; Younes, 1990, 2015). The second rationale pertains to the status of MSA and CA, where both varieties have many lexical and grammatical features in common. Therefore, introducing the two varieties from the outset would not increase the learning burden on Arabic L2ers (e.g., Al-Batal, 1992, 2018; Younes, 2015). The third rationale stems from many survey-based studies that have found a positive attitude held by Arabic L2ers toward learning two varieties from the outset of the learning process (e.g., Al-Batal & Glakes, 2018; Belanp, 2006; Hashem-Aramouni, 2011; Huntley, 2018; Husseinali, 2006; Isleem, 2018; Kuntz & Benlap, 2001; Palmer, 2007, 2008; Shiri, 2013).

However, from SLA perspectives, opponents of the IA consider it problematic for three main reasons. First, advocates of this approach appear to overlook at least four “guiding principles” that SLA researchers view as noncontroversial: clarity of input, frequency of input, integration of form and function, and integration of all language skills (Alhawary, 2013). Second, exposure to two linguistic forms conveying a similar meaning would highly likely cause confusion and result in blending the two forms by L2ers, who may thereby end up with no mastery of either form (Alhawary, 2009a, 2013; see also Nation & Macalister, 2010; Schwartz & Causarano, 2007). Despite this argument, proponents of the IA do not provide any practical solutions to the linguistic dissonance (i.e., confusion) that usually results from input optionality

(Alhawary, 2021; Eisele, 2018). Third, notwithstanding the shared ancestry and great similarities between MSA and CA, they exhibit major differences across all language domains—potentially increasing the learning burden on Arabic L2ers even more (e.g., Alhawary, 2013; Eisele, 2018). Based on these arguments, opponents of the IA believe that teaching *only* MSA *then* CA may be the optimal option to avoid any potential failure that might be the ultimate outcome of the simultaneous exposure to two varieties.

However, the debate between those in favor of teaching MSA-then-CA and those advocating for the simultaneous teaching of MSA and CA has remained unresolved. One key reason for this ongoing dispute is attributed to the lack of SLA-based studies that tackle this issue directly and that could provide empirical evidence supporting either stance (Alhawary, 2018; Wahba, 2023). Very few Arabic L2 studies have specifically examined this issue (Leddy-Cecere, 2018; Nassif, 2018; Shiri & Joukhadar, 2018). Although they found the IA to be beneficial to Arabic L2ers in terms of developing linguistic and sociolinguistic competencies similar to those of native speakers, these studies focused primarily on the code-switching behavior in the L2ers' language output. Additionally, these studies suffered from many methodological limitations, including the heavy reliance on rehearsed L2ers' output (e.g., presentations, skits) and the lack of analysis tracing the development of Arabic structures in the L2ers' interlanguage systems. Therefore, further SLA-based research that controls for relevant variables is much needed to shed more light on this issue.

## **1.2 Importance and Focus of the Dissertation**

The importance of the current study stems from the absence of empirical evidence that could be in support of the stance of either camp of instructional practices: teaching MSA-then-CA versus teaching MSA and CA simultaneously. As highlighted by Alhawary (2018, 2021) and Wahba

(2023), to ascertain which instructional approach yields superior learning outcomes, empirically data-driven studies on the acquisition of Arabic morphosyntactic structures by two types of Arabic L2ers are needed. Thus, this dissertation aims to fill in this gap by providing data-driven accounts of the learning outcomes of both pedagogical approaches by comparing the language output of 57 Arabic L2ers, representing two comparable groups: the first was instructed in MSA and EA simultaneously from the outset of the learning process, the integrated program group (IPG), and the second was instructed in MSA only, the unintegrated program group (UIPG). The selection of EA over other CA varieties was motivated by Egypt's desirability as a travel destination for Arabic L2ers in U.S. universities (Shiri, 2013) and the prevalence of EA as a commonly chosen dialect when teaching CA to Arabic L2ers in most U.S. universities (e.g., Eisele, 2018; MLA, 2016; Looney & Lusin, 2019). Moreover, a control group (CG) consisting of nine native speakers of EA was recruited to establish a baseline along which the performance of both L2 groups would be compared.

Many factors were considered in the present study. The first pertains to the variables that were controlled for in the recruitment process of the L2 participants of both L2 groups (IPG and UIPG). To control for the cross-linguistic effect variable, only English L1 adult speakers with Arabic as their L2 were recruited—deliberately excluding heritage speakers of Arabic. Additionally, to examine the development of their acquisition over time, the input exposure variable was strictly controlled for by categorizing the recruited L2 participants of each L2 group into three proficiency levels based on their formal exposure to Arabic: first year, second year, and third year.<sup>1</sup> Notably, both L2 groups were to a great extent comparable in terms of their prior

---

<sup>1</sup> The term *proficiency level* is loosely used here to indicate the extent of competence attained through the years of formal Arabic classes in which the L2ers were enrolled during the time of data collection.

exposure to Arabic at each proficiency level, and they exhibited statistical homogeneity in their self-perceptions of their Arabic proficiency and in their performance on the Arabic placement reading test administered to them before the data collection process took place.

The second factor considered in the present study revolves around the question of how to measure the L2ers' learning outcomes and development. Applied linguists and SLA scholars (e.g., R. Ellis & Barkhuizen, 2005; Housen et al., 2012; Housen & Kuiken, 2009; Norris & Ortega, 2009; Polat & Kim, 2014; Skehan, 2009; Unsworth, 2008; Wolve-Quintero et al., 1998) have reached a consensus on at least three dimensional constructs that can objectively and neutrally operate as predictors of the L2 overall proficiency and development. Those constructs are fluency (e.g., speed of output), accuracy (e.g., error-free output), and complexity (e.g., subordination production). Measuring the learning outcomes in terms of these three facets is crucial because they can "provide a sophisticated framework for investigating the multi-componential nature of language use and development" (Polat & Kim, 2014, p. 186). Thus, the language output of the participants was measured in terms of fluency, accuracy, and complexity, as this triad can meticulously reveal their unconsciously underlying competence and proficiency.

The third factor considered in the current study relates to the morphosyntactic target structures examined. The accuracy construct was examined in the acquisition of the structures of simple past/perfect tense/aspect, simple present/imperfect tense/aspect, past negation, and verbal present negation. As for the complexity construct, it was examined in the acquisition of the subjunctive subordination and cause-effect subordination. Three reasons underlie the choice of these structures over others. First, they are considered high-frequency structures in the input to which Arabic L2ers are exposed from early on. Thus, this can lay the basis for examining the role of input frequency in the development of these structures in the L2ers' interlanguage

systems. Second, MSA and EA diverge in these structures (except for the past/perfect tense/aspect), so examining such structures in the L2ers' language output may reveal the extent to which the IPG participants could comprehend these structures and master them in both varieties. Third, the examination of these structures, particularly those in which MSA and EA diverge, may reveal whether the simultaneous exposure to two varieties would result in confusion (and perhaps early fossilization) and hinder facilitative L1 transfer at an early stage of the learning process (Alhawary, 2013).

The fourth factor taken into account in the present work pertains to the instruments employed for data collection. The acquisition data were elicited in the form of one-on-one spontaneous elicitation sessions with all participants. Such data could realistically reflect the unconsciously underlying competence of Arabic L2ers across different stages of their interlanguage development (Alhawary 2009a, 2019). Moreover, picture-based description tasks were designed to ensure eliciting enough tokens of the target morphosyntactic structures in obligatory contexts only.

### **1.3 Objectives and Questions of the Dissertation**

As previously mentioned, the lack of consensus among Arabic L2 practitioners on how to approach the Arabic sociolinguistic status in L2 classroom settings has generated a relatively long debate that has not been resolved yet due to the absence of SLA-based studies tackling this issue directly. Therefore, the main objective of the current study is motivated by the need to contribute to the bulk of Arabic L2 literature by bridging the gap caused by the absence of empirical SLA accounts of the learning outcomes of the simultaneous exposure to two Arabic varieties from early on. In essence, examining and comparing the language output of the two groups of Arabic L2ers (IPG and UIPG) can provide solid evidence regarding which one of the

two instructional practices (teaching MSA-then-CA versus teaching MSA and CA simultaneously) is more practical and beneficial to Arabic L2ers in terms of enhancing their fluency, accuracy, and complexity constructs—which constitute their overall proficiency.

Furthermore, besides the attempt to relate the current findings to previous findings from the general SLA literature, the present work aims to account for the data in light of three factors that have been the center of the SLA literature and that contribute the most to the SLA process: the quality and quantity of input, L1 transfer, and implicit learning (i.e., incidental learning with focus on function) versus explicit learning (i.e., intentional learning with focus on form). Building on the complex dynamic systems theory (CDST), which perceives language as an innately complex process influenced dynamically by a variety of interrelated factors that drive the whole course of interlanguage development (de Bot et al., 2007; de Pot & Larsen-Freeman, 2011; Larsen-Freeman, 1997, 2017), the acquisition of some target structures examined here would be expected to be problematic for the L2 participants, depending on the internal and external resources available to them (e.g., input, L1 transfer, mode of learning).

To elaborate, since MSA, EA, and English converge in some of the target structures and diverge in others (see Section 3.3: Chapter 3; Section 5.2.2: Chapter 5), analyzing the acquisition data of the L2 participants in light of the notion of cross-linguistic effect and access to universal grammar (UG) can determine how the acquisition of these structures may be affected by the L1 transfer factor. However, in the Arabic L2 context, L1 transfer, as an internal mechanism, cannot solely account for the variations in the learning outcomes of L2ers because other factors appear to play influential roles as well. Those factors include the nature, quantity, and timing of input, as well as the mode of implicit versus explicit learning. Particularly, previous studies have found that negative/nonfacilitative L1 transfer can be minimized by the frequency of the target L2

forms/structures in the input (Alhawary, 2005, 2009a, 2009b, 2013, 2019; Azaz, 2016; Lin & Alhawary, 2018) and by the explicit (visual) presentation and instruction of them in the input (Al Midhwah & Alhawary, 2020). Moreover, negative/nonfacilitative L1 transfer may persist, and successful/facilitative L1 transfer may be hindered if the input is obscured, as predicted by various L1 transfer hypotheses, including the full transfer/full access hypothesis (FT/FAH; Schwartz & Sprouse, 1996), missing surface hypothesis (MSH; Lardiere 1998, 2000), feature reassembly hypothesis (FRH; Hwang & Lardiere 2013; Lardiere 2008, 2009), and robustness of L1 transfer hypothesis (RL1TH; Alhawary, 2021). Hence, interpreting the data in light of these factors will contribute to the ongoing discussion of the associated role of input, L1 transfer, and implicit versus explicit learning in the development of Arabic L2ers' proficiency.

Accordingly, this dissertation is guided by the following research questions (RQs):

**RQ1.** What is the nature of the language output of Arabic L2ers in the IPG versus that of Arabic L2ers in the UIPG across the three proficiency levels (first year, second year, and third year)? More specifically:

- Are there any similarities and differences between the IPG and UIPG in terms of their fluency when narrating past and present events?
- Are there any similarities and differences between the IPG and UIPG in terms of their accuracy on the target structures of past tense and present tense, as well as past negation and verbal present negation?
- Are there any similarities and differences between the IPG and UIPG in terms of their error-free production of the target complex structures of subjunctive subordination and cause–effect subordination?

- Are there any similarities and differences between the IPG and UIPG in terms of their overall proficiency?

**RQ2.** What is the role of input, L1 transfer, and implicit versus explicit learning in affecting the performance of the IPG versus that of the UIPG?

#### **1.4 Outline of the Dissertation**

This dissertation is organized into five chapters, including the current introduction chapter, which provides an overview of the study's motivation, importance, focus, objectives, and RQs. Chapter 2 offers a systematic literature review of the sociolinguistic (diglossic) status of Arabic, the pedagogical approaches adopted by Arabic L2 programs to diglossia, and the arguments for and against the IA. Then, Chapter 2 provides a comprehensive analysis of the three Arabic L2 studies on the simultaneous learning of MSA and CA and reviews the most widely accepted definitions and measures of the fluency, accuracy, and complexity constructs, as well as some representative Arabic L2 studies that have examined this proficiency triad. Chapter 2 concludes with a discussion of the factors of input, L1 transfer, and implicit versus explicit learning, along with their implications for the current study.

Chapter 3 provides a detailed discussion of the methodological instruments employed in the study, including the various criteria implemented in the recruitment process of the participants, followed by the recruitment results. The chapter also discusses the target structures, their presentation timing and frequency in the L2 participants' instructional textbooks, the tasks used, the data collection procedure, data coding, and data analysis.

Chapter 4 presents the findings revealed by the data elicited from the participants. The data were analyzed both qualitatively and quantitatively. The qualitative method entailed fine-grained delineation of the correct and incorrect rule application in the production of the target



structures. Thus, an ample amount of representative data samples were provided to offer a full scope of how the target structures were processed in the L2 participants' interlanguage systems across the three different stages of their L2 development. As for the quantitative analysis, it entailed provision of descriptive statistics and figures (boxplots and trending lines) for the mean scores of the IPG and UIPG at each proficiency level, as well as the CG. The analysis also entailed conducting statistical tests on the mean scores of the participants on fluency, accuracy, complexity, and overall proficiency.

Chapter 5 discusses the current findings in terms of the RQs and discusses their implications within the broader context of previous findings from the SLA literature. The chapter concludes by discussing the pedagogical implications of the study and by acknowledging its limitations.

## Chapter 2 Literature Review

### 2.1 Overview

Although the simultaneous teaching and learning of two Arabic varieties (i.e., MSA and CA) has been a pressing issue in the Arabic L2 field of study for at least the past decade, it has been overlooked and has not been investigated thoroughly from SLA perspectives. In fact, only three recent studies have examined the output of Arabic L2ers exposed to MSA and CA concurrently (Leddy-Cecere, 2018; Nassif, 2018; Shiri & Joukhadar, 2018). However, before delving into these studies, it is important to review the sociolinguistic status of Arabic to understand why the issue of teaching and learning two Arabic varieties emerged in the first place (Section 2.2). Subsequently, a synthesized discussion of the various approaches to Arabic as an L2 is offered, with a particular focus on the "integrated approach" (IA), along with the arguments supporting and opposing this approach (Section 2.3). Next, a detailed analysis of the three Arabic L2 studies on the simultaneous learning of MSA and CA is provided (Section 2.4). The chapter then proceeds with a discussion of the most the widely adopted definitions and measures of the three proficiency constructs of fluency, accuracy, and complexity, as well as some representative Arabic L2 studies that have examined this triad altogether (Section 2.5). Given that the data are discussed considering three key SLA factors—input, L1 transfer, and implicit versus explicit learning—these factors are thoroughly reviewed in Section 2.6. The chapter concludes with a summary of the reviewed literature and a discussion of the implications of the three SLA factors for the current study (Section 2.7).

## 2.2 The Sociolinguistic Status of Arabic

The dialectal variations and the coexistence of multiple Arabic varieties can be traced back as far as 15 centuries ago.<sup>2</sup> However, the dialectal variations were not put in a sociolinguistic conceptual framework until the second half of the 20th century, when Ferguson (1959a) discussed comprehensively and seminally the term *diglossia* by referring to four different linguistic situations: Arabic, Modern Greek, Swiss-German, and Haitian Creole.

According to Ferguson (1959a), diglossia refers to the parallel existence of two linguistic varieties within a speech community, namely the superposed variety and the regional dialect, with both belonging to the same language. The functional distribution between these two varieties is based on factors, including formality, prestige, acquisition versus learning, literary heritage, standardization, stability, grammatical complexity, vocabulary repertoire, and phonemic inventory. Ferguson assigned the label High (H) to the superposed variety (i.e., classical or standard) that is more formal, more prestigious, learned, used in literary heritage, standardized, more stable, and that exhibits complicated grammar, a richer repertoire of vocabulary, and an expanded phonemic inventory. Conversely, the status of Low (L) was assigned to the regional dialect (i.e., colloquial) whose usage is limited to informal and oral communication. To apply Ferguson's definition to Arabic, *al-fuṣḥā* "MSA" represents H, which is more formal, prestigious, standardized, and used in literary and media contexts across the Arab landscape. On the other hand, *al-ʿāmiyyah* "CA" represents L, which serves as a medium of informal and daily interactions and displays geographical variations across the Arab world.

---

<sup>2</sup> For elaborate descriptions and discussions of the dialectal variations that the different geographical regions of the Arabian peninsula exhibited 15 centuries ago, see for example, *al-Kitāb* (1988) by Sībawayh (d. 796), *Sirr Ṣināʿat al-ʿIṣṭihār* (2000) by Ibn Jinnī (d. 1002), *as-Ṣāhibī fī Fiqh al-Lughah al-ʿArabīyyah wa Masāʾilihā wa Sunan al-ʿArab fī Kalāmihā* (1997) by Ibn Fāris (d. 1004), *al-Muzhir fī ʿUlūm al-Lughah wa ʿAnwāʾihā* (1998) by Al-Suyūṭī (d. 1505), and *Fiqh al-Lughah wa Sirr al-ʿArabīyyah* (2002) by al-Thaʿālibī (d. 1471), and Wafi (2004).

However, Ferguson's (1959a) clear-cut binary distinction between MSA (i.e., H) and CA (i.e., L), whose usage is subject to their functional distribution, has been challenged by many Arabists and linguists who argue for the existence of a continuum of varieties that could represent the sociolinguistic situation of the Arabic-speaking world more accurately than Ferguson's binary conceptualization. For example, Blanc (1960) suggested five levels of Arabic that are frequently used by native speakers: standard classical, modified classical, semi-literary, koineized colloquial, and plain colloquial. Sequentially, Badawi (1973) adopted Blanc's (1960) proposal and described the Egyptian context, which constitutes one of the major speech communities in the Arab world, as exhibiting a continuum of five Arabic varieties that intermingle and overlap with one another in most daily conversations, regardless of the discourse context. Those varieties are *fushḥa t-turāḥ* "heritage Classical Arabic," *fushḥa l-ʿaṣr* "MSA," *ʿāmmiyyat al-muḥaqqafīn* "colloquial of the educated," *ʿāmmiyyat al-mutanawwirīn* "colloquial of the literate," and *ʿāmmiyyat al-ʿummiyyīn* "colloquial of the illiterate."

Subsequently, the notion of the variety continuum laid the basis for others (e.g., Mahmoud, 1986; Meiseles, 1980; Mitchell, 1986; Ryding, 1991) to formulate another concept, called Educated Spoken Arabic (ESA), to refer to a variety that consists of mixing structures of MSA and CA. Native speakers refer to this variety as *al-luḡa l-wuṣṭā* "middle language," signifying an intermediate mode of speech that cannot be strictly characterized as either *fushḥā* "MSA" or *ʿāmiyyah* "CA" (e.g., Mejdell, 2021; Versteegh, 2014).<sup>3</sup> In other words, the lexico-morpho-syntactic regularities of ESA cannot be attributed neatly to MSA or CA; rather, they come from both varieties.

---

<sup>3</sup> Other terms that have been used to refer to ESA include Formal Spoken Arabic, Third Language, Modern Language, Spoken Language (e.g., Al-Batal, 1992; Mejdell, 2021; Ryding, 1991, 2006). However, for consistency, the term ESA will be used henceforth.

Ryding (1991) asserted that ESA came into existence because of its prestigious shared features, mutual intelligibility, and expansion across the Arab world. Mahmoud (1986) likewise identified some crucial factors that contributed to its emergence, including the widespread use of education and media, as well as the prevalent use of technology. Another vital factor behind its emergence is attributed its extensive use by elites, including politicians, who opt for ESA as a mutually intelligible medium of communication to convey their enthusiastically ethno-nationalist speeches that call for a rejuvenation of Arab nationalism to all people of the Arabic-speaking world. Ryding (1991, p. 216) stated 10 salient features that are exhibited by ESA and that distinguish it from both MSA and CA. Most of those features are similar to those stated in Ferguson's Koine hypothesis (1959b). They include the following:

1. Omission of the final short vowels in all parts of speech
2. Consequent metathesis of vowels on pronoun suffixes
3. Reduction of inflectional endings in dual and masculine plural sounds to the oblique or nonnominative form
4. Elimination of the separate feminine plural categories in verbs and pronouns and reduction to one nongender-specific plural
5. Elimination of the dual category in verbs and pronouns and merging of this category with the plural
6. Omission of final *nūn* in inflectional suffixes for the second-person singular feminine and second-person and third-person plural in the imperfect
7. Generalization of the defective suffix stem to geminate verbs in the past tense
8. Reduction of the definite relative pronouns to one form: *ʔillī*
9. Conversion of final nunation in indefinite defective nouns to a long vowel
10. Creation of a category of verbs with embedded indirect object

The diglossic status of Arabic, or more precisely the existence of a continuum of Arabic varieties, resulted in adopting various pedagogical approaches that aim at equipping Arabic

L2ers with linguistic and sociolinguistic competencies that are native-like in nature. Those approaches are discussed immediately below.

### **2.3 Pedagogical Approaches to Arabic as an L2**

Since the mid-17th century, when the study of Arabic as an L2 started in the United States, the motive behind learning Arabic was due to purely theological, philological, and archeological purposes. Consequently, teaching Classical Arabic (i.e., the ancestor of MSA) based on the grammar-translation method (e.g., reading, grammar, and translation drills) was the optimal way to achieve these purposes (McCarus, 1987, 1992; Ryding, 2006, 2018). However, after World War II, the objectives of teaching L2s in the United States (and ubiquitously) shifted into “the audio-lingual mode which aimed at developing proficiency in speaking and listening as well as reading, writing, and translation. For Arabic, this posed a particular problem because of the issue of diglossia” (Ryding, 2018, p. 13). That is, while MSA (i.e., the H variety) is mainly and functionally used in literary across all Arabic-speaking societies, CA (i.e., the L variety) is predominately used as the medium of daily communication.

Despite the shift in focus toward a more comprehensive language skills approach, most Arabic L2 programs in the United States, according to Al-Batal (2018), have continued to prioritize the teaching of MSA only, as reflected in the course titles offered in most of those programs (e.g., Elementary MSA, Intermediate MSA). Al-Batal (1992, 2018) and Younes (2015) outlined a number of practical, historical, and ideological reasons that may have contributed to the continued dominance of the MSA-only pedagogical approach. First, due to the lack of consensus among Arabic teachers and program directors on which dialect(s) should be taught, they opt for the most convenient solution, which is teaching MSA only. Second, the prestigious status of MSA, as the language of *Qurʿrān*, the official language of Arab countries, and the

language used in education, has led many Arabic programs to prioritize its instruction over CA. Third, due to the limited class time, it is practically impossible to teach all dialects, leaving the MSA instruction as the only feasible choice. Fourth, because most literacy sources are written in MSA, teaching this variety is practically the only viable option available to Arabic L2 pedagogues. Fifth, some Arabs hold the belief that demands for teaching CA are rooted in imperialist and orientalist agendas and plans intended to distance Arabs from each other and from their cultural heritage. Such a belief has thus propelled them to strictly adhere to the MSA-only approach. Despite these factors and others, many Arabic L2 programs have started to pursue other pedagogical practices that incorporate a spoken form of Arabic alongside MSA—allowing L2ers to develop sociolinguistic and linguistic competencies similar to those of native speakers.

The most widely embraced approach is to initially teach MSA for the first 2–3 years, followed by a sequence of courses that focus solely on a particular dialect and that reflect the student preferences and faculty expertise. Subsequently, opportunities are offered for L2ers to study abroad in Arabic-speaking countries, where they can receive an adequate amount of authentic input in CA (Alhawary, 2013, 2021; Eisele, 2018; Ryding, 2006). However, this approach has been rejected due to the claim that it creates a “firewall separation vision” between MSA and CA, which does not “reflect the linguistic reality across the Arab world, where MSA and the dialects coexist harmoniously and interact and intersect constantly in a wide variety of spheres” (Al-Batal, 2018, p. 6, see also Al-Batal, 1992, 1995; Younes, 2015).

Ryding (1991) proposed an alternative approach to the MSA-then-CA approach when she called for incorporating ESA into the Arabic L2 curricula. Although she highlighted the notion that teaching MSA solely enhances proficiency, she maintained that it might result in productions that do not accurately reflect the linguistic reality of the Arab world. Therefore, she

suggested incorporating ESA into instruction because it seems to take place in semiformal discussions and conversations by a wide range of speakers of different Arabic dialects. This approach was adopted by some L2 institutions, such as the Foreign Language Institute of the U.S. Department, where foreign service officers showed interest in learning a supra-dialect that is neutral in communicating with Arab officials ubiquitously and that can bridge the gap between the formality of MSA for brief meetings and the informality of CA for professional discussions. However, Ryding stressed that ESA cannot replace MSA in academic programs, but it “can be used to supplement the traditional focus on literary Arabic” (p. 213). Nevertheless, ESA, as a pedagogically alternative approach, has not gained momentum in Arabic L2 programs because of the variability in the linguistic output in formal situations among native speakers with different regional dialects and because it is merely a communicative tactic employed by native speakers for contextual, experiential, and personal reasons (Alosh, 1997, as cited in Mejdell, 2021).

More recently, the “integrated approach” (IA) has emerged as an alternative to the three approaches discussed above: the MSA-only-approach, the MSA-then-CA approach, and the ESA approach. The IA calls for the simultaneous integration of MSA and CA from the outset of the teaching and learning of Arabic as an L2, and it grew momentarily across the United States.<sup>4</sup> Arguments supporting and opposing this approach are thoroughly discussed below.

### **2.3.1 Arguments for the IA**

Younes (1990, 2015) proposed the so-called IA and claimed that the simultaneous integration of MSA and CA from the outset is the “most logical, effective, and economical way to prepare the student of Arabic as a foreign language to deal successfully with the Arabic diglossic situation” (2015, p. 1). Drawing from his teaching experience of MSA and Palestinian Arabic (or Levantine

---

<sup>4</sup> For a discussion of some Arabic L2 programs adopting the IA at various U.S. universities, see Al-Batal (2018).



ESA) at Cornell University, Younes rejected the practice of teaching MSA exclusively or initially, claiming that it leads to two detrimental consequences for the learning process. First, it may frustrate contemporary Arabic L2ers, whose needs and goals for traveling to Arab countries and interacting with native speakers might not be fulfilled by focusing solely on MSA. Second, teaching L2ers how to convey daily-topic conversations (e.g., family, school, work) in MSA only may trigger “what Ryding (2006: 16) refers to as *Reverse Privileging*,” a situation in which the secondary variety (MSA) is taught, whereas the primary discourse (CA) is neglected. That is, L2ers are not provided with the opportunity to have meaningful communication with native speakers and are “shut out of any opportunity of reinforcement, which is essential for any language development” (Younes, 2015, pp. 29–30, see also Al-Batal, 2018; Huntley, 2018; Nassif, 2018; Shiri, 2013; for the same argument).

Moreover, Younes (2015) sustained that introducing the two varieties independently in separate tracks “involves a certain amount of duplication and waste, creates an artificial division in the language, and deprives the learner of the opportunity to develop the skill to navigate the two language varieties as native speakers do” (p. 38). This is so because “linguistic differences between *‘Āmiyyah* and *Fuṣḥa* are far outweighed by the similarities, and since most differences are predictable from one variety to the other” (p. 41). To this end, he urged Arabic L2 pedagogues and programs to embrace the IA by teaching MSA and CA simultaneously, without privileging one over the other. This instructional practice would reflect the reality of the Arabic-speaking world, where the two varieties (i.e., H and L) are in complementary distribution (see also Shiri, 2013), help L2ers communicate effectively in Arabic-speech communities, provide them with authentic input, and meet their needs and goals by taking the educated native speakers’ proficiency as a model: full command of listening in both MSA and CA, full command

of speaking in CA, full command of reading and writing in MSA, and full command of code-switching between MSA and CA.

Similarly, Al-Batal (1992, 1995, 2018) argued for the need of an alternative approach opposed to the exclusive or initial teaching of MSA. He claimed that limiting the instructional practices to MSA can create “a fake model of oral proficiency by presenting the students with an artificial variety that is not used by the native speakers since no one uses [formal Arabic] for daily-life situations” (1995, p. 123). Al-Batal (2018) stressed the importance of treating the various Arabic varieties as one concrete entity rather than separate and unrelated entities. Such treatment is required to reflect both the diglossic status and recently evolving reality of Arabic, where the widespread evolution and use of social media and technology have expanded the usage range of various CA varieties in previously MSA-dominant public spheres and have endorsed further interactions between MSA and CA. Additionally, treating Arabic as one indiscrete entity can contribute to the development of curricula that better align with contemporary approaches to language teaching and learning, such as the communicative approach, which aims at equipping L2ers with sociolinguistic competence alongside its linguistic counterpart. Therefore, Al-Batal advocated for the IA from the outset of Arabic L2 instruction by introducing MSA as a medium for written communication and a vehicle for spoken forms in certain formal situations and by introducing CA as the primary medium for oral communication. By adopting this approach, L2ers can develop a balanced proficiency in both MSA and CA, effectively navigate various linguistic registers and sociolinguistic situations in Arabic-speech communities, and gain an in-depth grasp of various aspects of the Arabic culture.

The most elaborate argument supporting the IA comes from various survey-based studies investigating the attitudes of Arabic L2ers toward such an approach (e.g., Al-Batal & Glakes,

2018; Belanp, 2006; Hashem-Aramouni, 2011; Huntley, 2018; Husseinali, 2006; Isleem, 2018; Kuntz & Benlap, 2001; Palmer, 2007, 2008; Shiri, 2013). These studies argue that Arabic L2 programs should teach MSA and CA simultaneously from early on because Arabic L2ers showed positive attitudes toward this pedagogical practice.

For example, Husseinali (2006) investigated the motivation of Arabic L2ers, particularly the orientations of heritage and nonheritage L2ers. Data were gathered through a survey administered to 120 L2ers studying Arabic in five classes at major U.S. universities. They were divided into two main groups: the heritage group, consisting of 50 heritage Arabic speakers and non-Arab Muslim L2ers, and the nonheritage group, consisting of 70 L2ers who came from various ethnic and religious backgrounds. The survey comprised 23 items divided into two main sections. The first section encompassed seven items focusing on demographic information, whereas the second section involved 16 items focusing on the reasons behind their Arabic study (travel, instrumental, and identification orientations). The results showed no significant differences between the two groups regarding the travel orientations, as most participants expressed their desire to converse with Arabs, particularly during their travels to Arabic-speaking countries. However, the nonheritage group differed significantly from the heritage group in terms of the instrumental orientations. The former was more motivated to learn Arabic in order to understand the political situations in the Middle East and to improve their employment prospects. On the other hand, the heritage group differed significantly from the nonheritage group with respect to the identification orientations. The heritage group showed stronger motivation to learn Arabic due to its strong association with Islam and Arabism.

Palmer (2007) likewise analyzed two surveys (in the form of multiple choice Likert questions) that were conducted by the National Middle East Language Resource Center to obtain

insights into the Arabic L2ers and teachers' perspectives on teaching and learning CA. The first survey involved 650 L2ers from 37 U.S. institutions, and the second survey involved 82 teachers at 30 U.S. institutions. The results revealed that the majority of L2ers expressed their eagerness to study Arabic for effective communication with native speakers (88%) and for traveling to Arab countries (80%). This result indicated to the author that L2ers were willing to learn spoken varieties that are used in daily conversations. However, most teachers (67%) reported that L2ers should receive instruction in MSA only, at least for the first 2 years, before shifting to CA.

Similarly, Shiri (2013) examined the reactions and attitudes of Arabic L2ers toward their experience of studying regional dialects in immersive Arabic programs in Arabic-speaking countries. Data were collected from an online post-program evaluation survey that focused on the attitudes of 371 L2ers who studied in different Arab countries, including Egypt, Tunisia, Jordan, and Oman. The participants had varying levels of Arabic proficiency, ranging from less than 1 year to over 3 years of exposure. The survey consisted of 21 items divided into the following subgroups: the desired host-countries, attitudes toward learning a dialect, attitudes toward learning more than one variety of Arabic, and the potential sources that may have affected the participants' attitudes. The analysis revealed several findings. First, Jordan then Egypt were the most desirable destinations for most participants. Second, 86% of the participants reported that it was extremely important to learn a dialect. However, only 25 participants emphasized that learning a dialect could enhance their communication with the local people and decrease their level of embarrassment. Third, 80% of the participants refused to learn MSA only, with 26 participants reporting that CA would aid in a more proper understanding of the culture and media. Fourth, 81% of the participants acknowledged the usefulness of learning more than one dialect, and 93% of the participants agreed that learning one dialect would facilitate the

acquisition of other dialects. Fifth, the potential factors that affected the participants' attitudes revolved around the importance of learning a dialect to travel to the Middle East and communicate with Arab friends.

Therefore, Shiri (2013) rejected the exclusive teaching of MSA and claimed that the simultaneous teaching of MSA and CA, regardless of which dialect(s), is the optimal solution for three reasons. First, L2ers who study abroad without knowledge of any regional dialect may struggle to communicate effectively with the host community. Second, those who do not have a chance to study abroad may finish their Arabic programs with no proficiency in any spoken variety. Third, their exclusive knowledge of MSA can be a source of embarrassment and frustration because of their use of a variety that is not spoken by the members of the host society—thereby making the acculturation process even more challenging. On this note, Palmer (2007) stated that “it is also not uncommon for native Arabs to snicker at foreigners who only speak the formal language, thus potentially causing a sense of humiliation” (p. 112; see also Hashem-Aramouni, 2011; Huntley, 2018; for a similar argument).

Al-Batal and Glakes (2018) also examined the perspectives and attitudes toward the IA among 184 Arabic L2ers at the University of Texas' Arabic program, which adopted the IA in 2009. Data were collected from a survey consisting of 33 multiple-choice items, some of which were in the form of leading questions (e.g., learning a dialect simultaneously with MSA is empowering). The survey items were categorized into eight groups, three of which were directly related to the participants' opinions about the integration, its linguistic benefits, and its nonlinguistic benefits. The results revealed that more than 90% of the participants felt that the IA was beneficial to them because it helped them develop their overall language proficiency, enhance their understanding of the Arabic culture, and build robust connections with Arabs.

However, while 94% of the participants supported the notion of integrating dialectal elements alongside MSA from early on, 66% of them reported that it is confusing.

The confusion issue was also echoed by Huntley (2018) in her analysis of a survey administered by the University of Michigan's Arabic program, which aimed at gathering the students' perspectives on the IA. Out of the 67 L2ers who completed the survey, 12 respondents "often felt confused about which register to use and when or otherwise lacked the ability to mentally differentiate between the two on the spot" (p. 79).

Thus, the above-reviewed representative studies endorsed Arabic L2 programs to correspond to the L2ers' demands and accommodate their needs by incorporating CA along with MSA in the Arabic curricula, which can help the L2ers fit in with Arabic societies and "avoid linguistic embarrassment" (Palmer, 2007, p. 120). This demand was put forward despite the high response rate pertaining to the teachers' preference of the exclusive use of MSA in the first 2 years of learning (e.g., Featherstone, 2018; Palmer, 2007) and despite the linguistic dissonance (i.e., confusion) that was highly reported by Arabic L2ers who were simultaneously exposed to MSA and CA from the outset (e.g., Al-Batal & Glakes, 2018; Huntley, 2018).

### **2.3.2 Counterarguments to the IA**

Opponents of the IA do not deny the advantages of the simultaneous teaching and learning of MSA and CA because such an instructional practice can provide Arabic L2ers with authentic input that reflects the language use in the Arab world and can also fulfill their needs and goals (Alhawary, 2021; Eisele, 2018). However, the disadvantages of this pedagogical approach outweigh the advantages. Criticism of the IA comes from SLA perspectives and is cognitively oriented, including the theoretical foundation on which it was developed, the potentiality of overwhelming Arabic L2ers with unclear and insufficiently recycled input, the linguistic

dissonance that might be the ultimate result of input optionality, and the real status of the similarities and differences between MSA and CA.

Any pedagogical approach must consider four essential issues that constitute its major components: understanding the nature of language, how an L2 is learned/acquired, the different methods and techniques that promote language learning/acquisition, and how to assess language learning outcomes (e.g., Richards & Rodgers, 2001). However, proponents of the IA do not discuss these essential components. On this note, Alhawary (2021) emphasized the following:

Although providing a sociolinguistic account is a necessary component for advocating for a particular pedagogical approach, three other components, which are at least equally necessary, are completely absent from Younes' treatment. No discussion is provided about the theoretical language learning underpinnings or how variation is assumed to be learned, given the variability or entailed optionality of language use (i.e., with such optionality a given code is not guaranteed to be learned or produced), timing of acquisition, and other input effect considerations. A second absent component which would be expected but completely missing is any discussion of a methodological rationale of how to test and assess learning, using the new approach. A third missing component is an explicit account of the linguistic input (i.e., where criteria for what, when, where, and how MSA is integrated with a colloquial or colloquials are not explicitly articulated) which is a pedagogical prerequisite for contents for drill and practice by the learner in the classroom and for language input modeling and error correction (by means of recast or other methods) by the teacher. (pp. 20–21)

The second counterargument to the IA pertains to the four guiding (input) principles perceived as noncontroversial by SLA researchers but overlooked and neglected by the proponents of this approach (Alhawary, 2013).<sup>5</sup> The first principle is the clarity of input principle, which suggests that L2ers can approximate further toward the target language if they receive transparent input that clearly represents the language's grammatical rules. If the input lacks clarity, L2ers will not understand it and thus might go through early fossilization. Therefore, input for at least first-year

---

<sup>5</sup> For a thorough discussion of the input factor in SLA, see Section 2.6.1 below.

and second-year L2ers must be transparent and realistic—that is, in one variety only. This can reduce the learning burden on L2ers whose “working knowledge” is very limited during the first 1–2 years of learning, which could result in incomprehension of the input if received in two or multiple varieties.

The second guiding principle is the frequency of input principle, which emphasizes the essential role of multiple exposure to *input* in converting it into *intake* in the L2ers’ interlanguage systems. Given that the American Foreign Service Institute (2020) ranks Arabic as a highly challenging language that requires 2,200 class hours from English L1 speakers to reach professional working proficiency, the 2–3 years of Arabic courses offered at universities (i.e., venues where learning Arabic takes place in foreign environments) can barely help L2ers develop “basic competence” that enables them to understand basic grammatical structures of Arabic (Abboud, 1968). Thus, it is not likely possible to sufficiently recycle the input in both varieties (MSA and CA) during these limited few years. Attempting to incorporate both varieties simultaneously within such a short time frame may lead to insufficient recycling of the input in both varieties, which can most likely lead to incomprehensibility of their structures.

The third guiding principle is the integration of form and function principle, which accentuates that focus on grammar is as essential as focus on other language skills. This notion is supported by the communicative approach itself, which seems to be misunderstood by many L2 practitioners who prioritize fluency and function at the expense of accuracy and form. In his discussion of the role of grammar instruction within the communicative approach, Thornbury (1999) affirmed the following:

And a glance at the so-called communicative course books confirm that grammar explanations are much more conspicuous now than they were, say, in the heyday of either the Direct Method or Audiolingualism . . . Without some attention to form, learners run a risk of fossilization . . . learning seems to be enhanced when



the learner's attention is directed to getting the forms right, and when the learner's attention is directed to features of the grammatical systems. (pp. 23–24)

Grammar instruction, especially explicit instruction, becomes more significant when teaching Arabic because it is a highly inflected language. However, since all Arabic dialects have not been thoroughly codified, CA grammar materials are mostly developed based on the personal experience of those adopting the IA to make up for the lack of descriptive accounts of the CA morphosyntactic structures. Consequently, explicit form-focused instruction is mostly neglected when teaching CA, resulting in inaccurate application of grammatical rules and unauthentic code-switching behavior between MSA and CA (Alhawary, 2013; Wahba, 2023).

The fourth guiding principle is the integration of language skills principle, which highlights that teaching language skills in isolation from one another is not effective. Language is best taught in an integrated manner to ensure reinforcement of all language skills, provide ample opportunities of input recycling, offer more practices on function and form, and sustain the L2ers' motivation throughout their learning journeys. However, because reading and writing are almost always carried out in MSA only, even in informal settings (as in Al Alaslaa & Alhawary, 2020, in the Saudi Twitter context), CA input would not be sufficiently and visually presented to L2ers.<sup>6</sup>

Another argument against the IA relates to the linguistic dissonance (i.e., confusion) that usually results from input optionality (Alhawary, 2021; Eisele, 2018). In fact, the findings reported in Al-Batal and Glakes' (2018) study provide evidence for this argument. As discussed earlier, 66% of 184 L2ers who participated in their study reported that it is confusing to be exposed to MSA and CA input simultaneously (see also Huntley, 2018). Despite some attempts

---

<sup>6</sup> Section 2.6 provides further discussions on the important role of input frequency, explicit learning, and visual presentation of the input in SLA development.

made by proponents of the IA to tackle this problem, they fell short of providing any practical solutions. Younes (2015), for example, argued that confusion is minimal in the integrated textbooks due to their design of separating MSA and CA. He further downplayed such an issue by stating the following:

The confusion argument might well be the result of an exaggerated concern on the part of teachers to protect their students from being overwhelmed, while students might in fact be more capable than their teachers give them credits for. If this is the case, then teachers would be doing their students a disservice by not preparing them for the sociolinguistic realities of Arabic while thinking they are helping them. (p. 56)

Likewise, Al-Batal and Glakes (2018) attempted to diminish the seriousness of confusion by claiming that it should not act as a barrier that prevents Arabic L2 programs from adopting the IA because confusion is a typical and natural phenomenon of any learning process and because it is also encountered by native speakers when they start learning MSA formally. They further advised Arabic L2 teachers to manage confusion by adopting the appropriate pedagogical method and by assuring their students that confusion is a normal part of the learning journey. On this note, Al-Batal (1992) stated the following:

Confusion that will be felt by the students should be regarded as part of the total experience of learning Arabic. This confusion is a reflection of native speakers of Arabic experience when they start their formal study of Arabic . . . Therefore, teachers should be prepared to deal with confusion and should make it clear to the students that the level of confusion will gradually diminish as they become more proficient in the language. (p. 302)

However, SLA scholars maintain that confusion should be taken seriously because it poses a huge problem for L2ers. According to Alhawary (2013, 2021) and Eisele (2018), L2ers hypothesize that they deal with one consistent set of linguistic features, but they start experiencing linguistic dissonance when they are provided at an early stage with alternative

options yet distinct structures that convey and represent the same meaning (i.e., input optionality). This concern is supported by Nation and Macalister (2010), who stressed that a language course should avoid the simultaneous teaching of linguistic items that are strongly related or similar to each other, either in form or function (e.g., opposites, synonyms). The simultaneous presentation of two strongly related forms can lead them to interfere with one another, which could potentially complicate the learning process even more. This was also echoed by Alhawary (2013), who argued this would amount to a learning burden on the L2ers. Likewise, Schwartz and Causarano (2007) assured that lumping optional but distinct structures might cognitively overwhelm L2ers and cause them “a great deal of confusion, resulting in production errors” (p. 53).

Empirical evidence in support of the confusion argument comes from one of Alhawary’s (2013) representative findings, where Arabic L2ers encountered difficulty in processing the MSA past negator *mā* “did not” because it was introduced with *lam* “did not”, as an optional negation particle, without sufficient recycling of the former (for a similar finding, see also Albirini et al., 2019; Alhawary, 2009a). Thus, Alhawary (2021) made a strong argument regarding this issue by affirming the following:

With the current cognitive constructionist paradigm of language learning, including the adoption of the communicative approach, adult L2 learners do not learn that which they do not understand. As is widely known by second language practitioners, this was one of the main causes for the abandonment of the audio-lingual approach and the adoption of the communicative approach. (p. 23)

The fourth argument against the IA relates to the similarities and differences between MSA and CA. While it is true that all Arabic varieties exhibit remarkable similarities, as argued by Al-Batal (1992, 1995, 2018) and Younes (2015), they also exhibit significant differences across all language domains. Those differences can have a negative impact on the whole learning process

if learned or taught simultaneously (Eisele, 2018). Even if those differences were assumed to have a minimal or no impact on the learning process, the mutual intelligibility between all Arabic varieties, as argued by Eisele (2018), would be in favor of teaching MSA initially. This is so because MSA serves as the language of literacy, contrary to CA, which is used mainly for oral communication purposes in informal settings.

Based on the aforementioned counterarguments, opponents of the IA believe that the simultaneous learning and teaching of two varieties from the outset can slow the learning process, result in overwhelming L2ers, and lead to confusion caused by the difficulty in comprehending and internalizing the input. As a consequence, problematic output and perhaps learning failure might be the ultimate outcome. Alternatively, Arabic L2ers should be exposed to “only one form at a time, and this is the logic of many in the field of Arabic language pedagogy who maintain that teaching MSA is the best way to proceed, at least for the first few years of instruction” (Eisele, 2018, p. 8), during which they can develop “basic competence” (Abboud, 1968) or “working knowledge” (Alhawary, 2013, 2021) that can lay a solid foundation for future learning of additional Arabic varieties.

#### **2.4 Studies on the Learning Outcomes of the IA**

Despite the lack of consensus among Arabic L2 researchers and pedagogues on the question pertaining to which one of the two common pedagogical practices (i.e., MSA-then-CA versus the IA) is more effective and efficient, only three studies have examined the production output of Arabic L2ers exposed to two Arabic varieties simultaneously (Shiri & Joukhadar, 2018; Leddy-Cecere, 2018; Nassif, 2018). Although these studies focused mainly on the code-switching behavior in the L2ers’ output, they are discussed extensively below due to their strong relevance to the current study.

Nassif (2018) aimed to examine Arabic L2ers' ability to simultaneously develop linguistic and sociolinguistic competencies in MSA and CA. Oral data were collected from 70 Arabic L2ers representing three proficiency levels: first year, second year, and third year. They were instructed in either MSA-EA or MSA-LA and used the third edition of *Al-Kitāb fī Ta'allum Al-'Arabiyya* (Brustad et al., 2011), which focuses on the integrated presentation of MSA and CA materials. Data from the first-year participants were collected through individual interviews about upcoming weekend activities. Data from the second-year participants were gathered through presentations (4–5 minutes), where they were instructed to use the most appropriate variety, depending on the presentation topics. Data from the third-year participants were collected from presentations (10–12 minutes) on formal topics that required the use of MSA more than LA or EA, as well as from skits (7–10 minutes), each of which involved a dialogue in LA. The data were qualitatively analyzed, focusing on verbs, negation particles, and conjunctions. The first-year participants' production data revealed a tendency to use CA features more than MSA, especially in negation markers and conjunctions. However, MSA features, as well as common features between MSA and CA, were present in their output. One notable observation was the use of CA-like prefixes with MSA-like verbs (e.g., the future aspectual marker *ḥa-* with the verb *nuqātil* "we will fight"). Nassif claimed that this usage should not be considered incorrect. Although it is prescriptively ungrammatical, this hybridization is attested in the speech of educated native speakers of Arabic (e.g., Gamal Abdel Nasser). The second-year participants' production data revealed more frequent usage of MSA features, along with common MSA-CA features, suggesting that the participants were developing more awareness of the appropriate variety that should be used in a given context. The language output of the third-year participants revealed more systematicity in terms of feature use. That is, MSA and common

MSA-CA features were used in the presentations, whereas CA features were predominant in the skits. Accordingly, Nassif concluded that the L2ers were able to develop diglossic competence resembling that of native speakers, who are conscious of which variety is expected to be used in specific contexts.

Similarly, Leddy-Cecere (2018) attempted to provide a qualitative account of the code-switching patterns observed in the language output of Arabic L2ers across four proficiency levels: first year, second year, third year, and fourth year. Data were gathered from field notes taken during 24.5 hours of Arabic class observations at a U.S. university, as well as from interviews with 16 L2ers at the same university. The interview questions were about descriptive topics (e.g., what is your favorite hobby?) and abstract topics (e.g., is the study of foreign language an important part of education?). It was assumed that these types of questions would reveal some code-switching or mixing patterns between MSA and CA in the L2ers' language output. The data revealed four general stages of code-switching. In the first stage, the code-switching behavior was limited in the output of the first-year participants who exhibited a tendency to use one Arabic variety in their discourse. As for the type, most code-switching patterns in this stage were in the form of alternation, insertion (e.g., *minšān huwwe la yurīd ʔan yaʔkul* “**because he** does not want to eat”), and repair/recast (e.g., *ʔalašān . . . liʔnna ʔal-wālida māta* “**because . . . because** the father died”). The limited code-switching behavior was motivated by the first-year L2ers' cognitive tendency to make their production more homogenous by sticking to one form of speech. In the second stage, the code-switching behavior increased among the second-year participants—although MSA remained the dominant variety used. The main type of code-switching observed in this stage was in the form of congruent lexicalization (e.g., *huwa yajlis fī ʕ-šālūn wa yiʕmal šī yarīb ʕašān hēk* “he is sitting in the living

room and doing an odd **thing because of this**"). The third stage showed systematic code-switching behavior in the output of third-year participants. Although MSA still dominated their output, some participants demonstrated "a predictable distribution of MSA versus other varieties on the basis of topic and stance" (p. 210). The prime types of code-switching in this stage were similar to those in the second stage, with the exception that the participants in the third stage started leveling out the marked forms in a dialect and substituting them by the unmarked forms in two or more dialects (e.g., /g/ is substituted by /j/, word-final /a/ is preferred over the /e/, the plural marker *-ūn* is replaced by *-ū*). In the fourth stage, the code-switching behavior exhibited a decreasing tendency in the output of the fourth-year participants. However, certain types of code-switching behavior (e.g., congruent lexicalization, alternation, repair, and insertion) were still attested. What distinguished this stage from the third stage was that MSA was no longer dominant in the output, but it exhibited rather a diglossic relationship with the other dialects, indicating that the L2ers' output resembled that of native speakers.

The third study was conducted by Shiri and Joukhadar (2018), who aimed to examine the L2ers' attainment in speaking MSA and CA. Data were collected from video recordings of interactions between two instructors and 36 first-year L2ers, as well as from skits and presentations. The Arabic course, in which the L2 participants were enrolled, was divided into two parallel sections. One section used MSA materials from Part I of the third edition of *Al-Kitāb fi Ta'allum Al-'Arabiyya* (Brustad et al., 2011), whereas the other section used dialectal (LA and EA) materials, covering various novice-intermediate topics (e.g., family, shopping, greetings), as well as authentic materials (e.g., songs, film clips). The L2ers were not permitted to use MSA in the dialect section and vice versa. Any instances of using the wrong variety were promptly and explicitly corrected by the instructors, who provided the L2ers with the equivalent

form from the appropriate variety. The results revealed more mixing in the dialect section, compared to the MSA section. Regarding the L2ers' awareness of mixing, 15 self-corrections to the right variety took place. As for the types of mixing, they primarily occurred at the word level, phonological level, morphological level (mainly in verb conjugation), and grammatical level (mostly in negation). The data from the skits and oral presentations revealed that the L2 participants were capable of producing complete sentences, and their output was mostly in the variety designated for the section. However, the length of output in the class interactions was shorter than its counterpart in the skits and presentations. To minimize code-switching in the class and maximize the L2ers' awareness of the differences between MSA and CA, the instructors employed three main strategies: direct and indirect corrective feedback, recasting, and modeling. Based on these results, the authors concluded that the simultaneous yet distinct teaching of MSA and CA helped the L2ers approximate further to the intermediate-level characteristics specified by the American Council on the Teaching of Foreign Languages (ACTFL, 2012). In addition, the L2ers developed awareness about the differences between MSA and CA, reached a high level of accuracy in their production, and rarely used any hybrid form that would render their use of both varieties unnatural or unauthentic.

The above-reviewed studies suffer from many methodological limitations and flaws. First, the studies lacked an adequate number of samples from the participants' output, which makes it challenging to validate the authors' conclusions. Of all three studies, only Nassif (2018) provided a few samples of the L2ers' production. However, as highlighted by Alhawary (2021), those samples were "neither natural nor authentic" (p. 22). In fact, the mixing elements from MSA and CA observed in the samples support the linguistic dissonance experienced by the L2ers, as they were likely confused about which elements belong to CA and which belong to



MSA. Second, the authors did not elaborate on the instruments employed for data collection, nor did they provide enough information about the participants' backgrounds (e.g., whether they were heritage speakers, exposure to Arabic before joining their current programs).

Third, in Shiri and Joukhadar's (2018) study, the participants were at the beginner level, suggesting that most of their language production was likely at the phrase level and about very simple topics. Therefore, those limited and novice productions would not provide enough evidence of confusion, fossilization, and even mixing. This is supported by one of Leddy-Cecere's (2018) findings, where code-switching between MSA and CA was limited among first-year L2ers, who generally tended to use one variety.

Fourth, the tasks employed in these studies were not appropriate due to their lack of spontaneity, which is considered a crucial criterion for measuring competence development (Alhawary, 2009a, 2019, 2021) and implicit knowledge (R. Ellis, 2009b). To illustrate, Nassif (2018), as well as Shiri and Joukhadar (2018), relied heavily on presentations and dialogue skits to collect their data, which could have hindered the L2ers' spontaneity because they rehearsed and perhaps memorized by rote the content of their presentations and skits. This is supported by Shiri and Joukhadar's (2018) findings, where the errors rate in class interactions averaged 9%, compared to approximately 5% in the skits and presentations—although the length of production was shorter in the class interactions than in the skits and presentations.

Fifth, even though Nassif (2018) stated that "the data also show examples of word-level hybridization along the same lines of the linguistic choices that Arabic speakers make" (p.181), neither Nassif nor Leddy-Cecere (2018) or Shiri and Joukhadar (2018) relied on a conceptual framework, such as the ESA common features stated in Ryding's (1991) study, to examine whether the L2ers' output aligns with the criteria of speech produced by educated native

speakers. Furthermore, none of these studies employed a control group whose performance could have functioned as a baseline along which the L2ers' production could have been compared.

Lastly, of all three studies, only Nassif (2018) focused on specific linguistic features. This contrasts with Leddy-Cecere's (2018) claim that his findings provide counterevidence to the potential fossilization that results from exposing L2ers to multi-varieties from early. However, he did not focus on specific features, nor did he trace their development in the L2ers' output across different stages "to prove or disprove existence of fossilization" (Alhawary, 2021, p. 22).

Given the scarcity of studies examining the simultaneous acquisition of MSA and CA from early on and considering the methodological limitations of the available very few studies, it is imperative to examine the efficacy of the IA. This examination entails comparing and analyzing the spontaneous language output of Arabic L2ers who are exposed to two varieties versus those who are exposed to one variety in light of three strands of significantly contributing factors to SLA development: input, L1 transfer, and implicit versus explicit learning. However, prior to delving into these factors, it is essential to review the three proficiency constructs—namely fluency, accuracy, and complexity—that are examined in this study.

## **2.5 Measuring L2 Development**

A fundamental issue to applied linguists and SLA researchers relates to the question of how to provide a clear and accurate picture of L2ers' learning development. The seriousness of this question lies in the fact that most institutional and commercial sectors in most speech communities demand accurate and rigorous assessments of L2ers' performance. These assessments can significantly influence important decisions pertaining to educational and employment issues, such as hiring, accreditation, and salary-related matters (Bachman & Clark, 1987). Thus, various assessment measures have been employed, including impressionistic

judgment, institutional status, in-house assessment instruments, and standardized proficiency tests have been used (Unsworth, 2008). However, these measures come with certain limitations. For instance, standardized proficiency tests focus primarily on the functional and communicative aspects of language use (Bachman & Clark, 1987). Consequently, testers' judgment regarding the extent to which L2ers' production fulfills these notions often involves subjectivity. Due to this and other limitations, applied linguists and SLA scholars have reached a general consensus on at least three main facets that can be used as predictors of the L2 proficiency and development: fluency, accuracy, and complexity (e.g., R. Ellis & Barkhuizen, 2005; Housen et al., 2012; Housen & Kuiken, 2009; Norris & Ortega, 2009; Polat & Kim, 2014; Skehan, 2009; Skehan & Foster, 2012; Unsworth, 2008; Wolve-Quintero et al., 1998). These components are also reflected in the ACTFL Proficiency Guidelines (2012), which suggest that language users become more fluent and produce more accurate and complex language as they progress in their proficiency—from the novice to intermediate to advanced to superior to distinguished levels.

However, SLA researchers disagree on what is meant by fluency, accuracy, and complexity, as various definitions of each construct coexist. This disagreement has been problematic for researchers, who are left undecided about which definitions to adopt. Housen et al. (2012) acknowledged this issue and affirmed that L2 studies “do not explicitly define what they mean by these terms, or when they do, they do so in rather general and vague terms . . . . As a result, terms *fluency*, *accuracy* and *complexity* are often used with different meanings across studies” (p. 4). The lack of consensus on what is meant by each construct has led to adopting various measures, consequently yielding inconsistent results across SLA studies (e.g., Housen et al., 2012; Norris & Ortega, 2009; R. Ellis & Barkhuizen, 2005). The immediate sections below provide a review of the concepts of fluency, accuracy, and complexity, along with their

respective measurement methods, followed by a discussion of several representative Arabic studies examining these three constructs all at once.

### **2.5.1 Defining and measuring fluency**

Housen et al. (2012) defined fluency as the L2ers' "global language proficiency, particularly as characterized in terms of the ease, eloquence, smoothness, and native-likeness of speech or writing" (p. 4). Fluency usually takes place when L2ers prioritize focus on function over focus on form (R. Ellis & Barkhuizen, 2005) and implement communicative strategies, such as circumlocution, paraphrasing, and illustration (ACTFL, 2012; Fakhri, 1984).

The bulk of SLA literature stresses that fluency is not a unitary concept and can be further divided into at least three subcomponents: speed fluency (i.e., speed of output), repair fluency (i.e., reformulation, false starts, and repetitions), and breakdown fluency (i.e., silence; R. Ellis & Barkhuizen, 2005; Housen & Kuiken, 2009; Skehan, 2009; Wolve-Quintero et al., 1998). Therefore, many methods have been proposed to measure this proficiency construct. One method suggests gauging fluency by computing the average number of words in produced sentences (e.g., R. Ellis & Barkhuizen, 2005; Wolve-Quintero et al., 1998).

Another method involves gauging fluency based on the mean length of pause (MLP). Vercellotti (2012) argued that measuring the MLP is equally important as measuring the speech rate because both pauses and speech rate have a mutually influential correlation. In general, longer pauses are expected to increase when the speed rate decreases and vice versa. However, some SLA studies, as noted by Ginther et al. (2010), have found no correlation between pauses and fluency. A third method proposes measuring fluency based on the principle of how fast L2ers access and produce linguistic units (syllables, words, phrases, clauses, or sentences) during a specific unit of time (second, minute, or conversational session; e.g., Ellis & Barkhuizen, 2005;

Ginther et al., 2010; Kormos, 2014; Skehan & Foster, 2012; Vercellotti, 2012; Wolve-Quintero et al., 1998). This method has been the most embraced method in Arabic L2 studies that have examined fluency, as will be discussed in Section 2.5.4 below.

### **2.5.2 Defining and measuring accuracy**

Housen et al. (2012) defined accuracy as “the extent to which an L2 learner’s performance (and the L2 system that underlies this performance) deviates from a norm (i.e., usually the native speaker)” (p. 4). According to this definition, the accuracy of L2ers is assessed based on their ability to produce error-free output that is native-like in nature. Thus, prescriptively clear and detailed rubrics outlining grammatical and ungrammatical structures should be provided to those who intend to gauge L2ers’ accuracy. However, some standardized proficiency tests, including ACTFL, have obscure and vague guidelines that can leave testers on their own in regard to figuring out what constitutes basic versus complex or correct versus incorrect language output—and hence misidentify the exact level of L2ers (Alhawary, 2021). In fact, this issue raises a serious problem for those who aim to measure the accuracy of Arabic L2ers in CA. Given that all CA varieties have not been thoroughly codified (Wahba, 2023), L2ers’ accuracy in CA may be misgauged by language testers and researchers due to the unavailability of detailed guidelines of what constitutes accurate structures in one CA variety but not in others and so on.

Another issue pertaining to accuracy is that there is no correlation between automaticity and accuracy. In other words, accuracy does not always indicate fluency and vice versa. On this note, Wolve-Quintero et al. (1998) maintained that “in second language learning, if the representation is faulty but access is fully automatized, it may lead to production that is error-full but impervious to change (perhaps the source of what has been called fossilization)” (p. 33).

Accuracy is often measured holistically by means of quantitative calculation of errors produced per unit of time (e.g., minute) or per linguistic unit (e.g., clause, T-unit). Typically, T-unit refers to the smallest language sentence that cannot be further split into grammatically meaningful sentences. For example, the Arabic sentence *huwa šariba ʔal-māʔ* “he drank water” is considered a single T-unit, but the sentence *huwa šariba ʔal-māʔ wa ʔakala ʔat-ʔafām* “he drank water and ate food” is considered one sentence consisting of two T-units (as in Albirini, 2018, 2019; Raish, 2018). The holistic error-free measure gained widespread popularity among SLA researchers because it was based on the principle that “the more developed/advanced L2ers are, the less likely they are to commit errors” (Unsworth, 2008, p. 319). However, as mentioned previously, before conducting any study investigating L2ers’ proficiency, it is vital to establish criteria for defining errors. Thorough descriptions of all L2 structures under study are therefore essential prerequisites (see Sections 3.3 and 3.6: Chapter 3).

### **2.5.3 Defining and measuring complexity**

Complexity is widely acknowledged as the most difficult, nebulous, and least comprehended concept of the proficiency triad. Consequently, a universally accepted definition of L2 complexity does not exist among SLA researchers. This lack of a precise definition has resulted in inconsistencies and sometimes contradictions in findings across SLA studies that have investigated L2 complexity (Bulté & Housen, 2012).

Some SLA scholars have maintained that complexity is generally investigated from a syntactic point of view only. They perceived L2ers’ complexity as “elaborated language” if (a) they produce language that is above the limit of their current interlanguage systems (i.e., language output that has not fully been internalized), and (b) they are willing to take risks by using a wide range of different complex structures (R. Ellis & Barkhuizen, 2005; Norris &

Ortega, 2009). However, this definition poses some challenges because it assumes that only language productions that have not been fully internalized can be considered complex, which disregards the fully automatized complex language. In other words, this definition excludes complex language if it is produced effortlessly and automatically.

Other SLA researchers have argued that complexity should not be interpreted from a syntactic point of view only but should encompass all linguistic components, including lexicon. For example, Wolfe-Quintero et al. (1998) perceived complexity as “a wide variety of both basic and sophisticated structures and words that are available to the learner” (p. 69). However, the inclusion of lexicon as a predictor of complexity is problematic due to the difficulty in deciding which lexemes should be considered complex and which should not.

A third group of scholars has provided a more rigorous definition of complexity. Bulté and Housen (2012), for example, argued that L2 complexity can be conceptualized as a construct consisting of two strands: linguistic complexity and cognitive complexity. While the former relates to the L2 features, the latter relates to the L2ers’ cognition. Linguistic complexity refers to “the degree of elaboration, the size, breadth, width, or richness of the learner’s L2 system or repertoire” (p. 25). On the other hand, cognitive complexity refers to the difficulty that L2ers encounter when processing linguistic features whose internalization is controlled by external variables (e.g., input frequency) and internal variables. The latter set of variables, as suggested by Robison (2001), can be further divided into affective variables (e.g., motivation, attitude, confidence, motivation, anxiety) and ability variables (e.g., intelligence, aptitude, and cross-linguistic effects). Apparently, this definition offers a clear explanation of complexity because it considers not only the complex nature of structures but also emphasizes the need for defining the

complexity of any linguistic feature in relation to the L2ers' external variables (e.g., input) and internal variables (e.g., cross-linguistic effects).

Various types of measures have been proposed as indicators of syntactic complexity. One of these measures relies on the notion of length of a produced unit, such as words produced per T-unit (W/T), which is commonly used in the field of first language acquisition (FLA; R. Ellis & Barkhuizen, 2005; Norris & Ortega, 2009; Wolfe-Quintero et al., 1998). Another widely embraced method in L2 studies involves measuring syntactic complexity based on the notion of subordination or complex T-unit, which is generally defined as one main clause containing at least one subordinate clause attached to or embedded within the matrix clause, such as relative clauses and cause-effect subordination (e.g., R. Ellis & Barkhuizen, 2005; Norris & Ortega, 2009; Wolfe-Quintero et al., 1998). Reliance on subordination as a measure of complexity is a valuable tool because “psycholinguistic studies have found that certain embedded structures (e.g., relative clauses) and passives are harder to process, or emerge later in language acquisition, than other structures” (Bulté & Housen, 2012, p. 21).

#### **2.5.4 Arabic studies on fluency, accuracy, and complexity**

In the Arabic L2 context, a few studies have examined the three constructs of fluency, accuracy, and complexity altogether. For example, Albirini (2018) investigated the proficiency of 29 beginner Arabic heritage speakers learning MSA as an L2 in relation to their socio-effective, social-contextual, and demographic variables. A control group consisting of 20 native speakers of Arabic was recruited as well. The participants completed two oral production tasks (one in MSA and one in CA) and a written task (in MSA), as well as a questionnaire focusing on various factors influencing proficiency, such as language input, language attitudes, identity, family pressure, community support, Sunday school, and religious practice. The proficiency of the



recruited groups was measured based on the three constructs of fluency, accuracy, and syntactic complexity. Fluency was measured by counting words produced per minute (W/M). Accuracy was measured by computing the ratio of error-free T-units produced to the total number of T-units produced. Complexity was measured by calculating the ratio of complex T-units produced to the total number of T-units produced. The overall proficiency of the two groups was measured by adding up the scores obtained in these three proficiency facets. The results revealed that the heritage group was significantly less fluent, accurate, complex, and proficient in speaking and writing compared to the control group. The significant difference between the two groups was attributed to differences between them in terms of the acquisition time of MSA and amount of exposure to it. Another finding pertains to the correlation between the proficiency in CA (i.e., the L1 of the heritage group), socio-affective factors (e.g., attitudes, identity, religious practice), external factors (e.g., input exposure to MSA), and socio-contextual factors (e.g., family pressure, community support) on the heritage speakers' proficiency in MSA. Of all factors analyzed, proficiency in CA (i.e., internal factor) and exposure to MSA input (i.e., external factor) were the only two predictors that could explain the variability in their proficiency in MSA. That is, those who received more input in MSA and those who were proficient in CA demonstrated higher proficiency in MSA, compared to those with little MSA input and low proficiency in CA.

Similarly, Albirini (2019) employed the same measures of fluency, accuracy, complexity, and overall proficiency when he attempted to address the question of whether MSA should be considered an L1 or L2 for native speakers of Arabic. Data were collected from 16 educated native Arabic speakers who learned English as an L2. Their performance in English was compared to that of 10 native speakers of English. The Arabic L1 group completed three oral

tasks (one in MSA, one in CA, and one in English) and three writing tasks (one in MSA, one in CA, and one in English). The English L1 group completed only one oral task and one writing task in English. The results were divided into three sections. The first section was related to the participants' oral proficiency in MSA, CA, and English. The data from the first section revealed a number of findings. First, the L1 Arabic group produced fewer words in MSA (75.8 W/M) than they did in CA (97.5 W/M) and in English (81.5 W/M), whereas the English L1 group produced the highest number of words (116.9 W/M). Second, the accuracy rates of the Arabic L1 group were comparable in MSA and CA (98.3% and 96.4%, respectively). However, their accuracy in English differed significantly from that of the English L1 group (that is, 75% versus 100%, respectively). Third, the Arabic L1 group produced the highest number of syntactically complex T-units in CA (28.3%) then in MSA (21.7%) then in English (19.8%). However, the English L1 group produced 30% complex T-units of all T-units produced. The second section of results focused on the participants' fluency, accuracy, and complexity when writing in MSA, CA, and English. The data of the second section yielded a number of findings. First, the Arabic L1 group produced comparable W/M words in MSA (20.1 W/M), CA (19.8 W/M), and English (19.9 W/M), whereas the English L1 group produced 27.3 W/M. Second, the Arabic L1 group was more accurate in MSA (97.3%) then in CA (87.2%) then in English (70.1%). As for the accuracy rate of the English L1 group, it was at ceiling (100%). Third, the Arabic L1 group produced more complex T-units in MSA (45.3%) than they did in CA (30.8%) and English (32.1%). As for the English L1 group, the rate of complex T-units was 49.1%. Based on the findings of the first and two sections, the English L1 participants significantly outperformed their Arabic L1 counterparts in terms of their oral and writing proficiency in English. Regarding the third section of findings, it focused on the oral and writing fluency, accuracy, complexity, and

overall proficiency of the Arabic L1 group in MSA, CA, and English. The statistical analysis revealed that their proficiency was comparable in MSA and CA. However, their proficiency in these two Arabic varieties was significantly better than its counterpart in English. Thus, Albirini concluded that MSA should not be considered an L2 for native speakers of Arabic due to their comparable fluency, accuracy, complexity, and overall proficiency in MSA and CA.

Raish (2017) also examined the fluency, accuracy, and complexity of Arabic L2ers in writing. He analyzed the writing of 14 English L1 advanced Arabic L2ers in comparison to that of 14 Arabic native speakers. The participants were asked to compose three essays targeting three genres: description, narration, and persuasion. He measured fluency by counting words produced per minute (W/M.), accuracy by calculating the ratio of syntactic error-free T-units produced to the total number of T-units produced, and complexity by counting words produced per T-unit (W/T). The results showed the native speakers were more fluent and accurate writers than the L2ers. However, the L2ers were as complex writers as the native speakers, especially in the persuasion-genre task, because both groups produced a comparable amount of W/T.

Notwithstanding the significant contribution of the above-reviewed studies, they exhibit many limitations related to the measures adopted. The accuracy construct was measured holistically in these studies (i.e., quantitative calculation of error-free T-units produced to the total number of T-units produced). However, exclusive reliance on this macro-based measure does not provide qualitative accounts of the types of errors produced. Consequently, it is unlikely to detect the causes for such errors, and hence it is difficult to obtain more insight into the development of the L2 among the L2ers, as suggested by R. Ellis and Barkhuizen (2005) and Wolfe-Quintero et al. (1998). In addition, Albirini (2018, 2019) measured complexity by computing the ratio of complex T-units produced to all T-units produced, but he again did not

provide any qualitative analysis of which of the complex T-units were correct and which were not. Furthermore, Raish (2018) measured complexity by counting words produced per T-unit (W/T). This measure may not determine whether the participants were complex writers because L2ers generally employ strategies (e.g., memorization of formulaic expressions) that can lengthen and expand their language output (R. Ellis & Barkhuizen, 2005; Vercellotti, 2012).

The current study attempts to avoid the above-discussed limitations by measuring the accuracy and complexity constructs at a micro level. Specifically, it examines accuracy in the production of four simple target structures (i.e., simple T-units) and complexity in the production of two complex target subordination structures (i.e., complex T-units) in obligatory contexts only (see Sections 3.3 and 3.4: Chapter 3). Furthermore, besides the quantitative analysis of the data, the present study implements a qualitative method to provide a fine-grained analysis of the correct and incorrect rule application in the production of the target structures.

Although some may argue that the examination of a few L2 structures does not provide a full picture of the L2ers' overall command of the language (R. Ellis & Barkhuizen, 2005), it is crucial to note that provision of a descriptive analysis of the L2ers' production in all language structures is not likely possible unless a longitudinal study is conducted on a very small pool of L2ers. Thus, the optimal alternative is to use the micro-based measure by tracing the development of particular L2 features in L2ers' interlanguage systems. This method has been the most embraced method across SLA studies, especially in cross-sectional studies.

Another argument that can be raised against the micro-based measure is that examining several L2 features only might provide a deceptive picture of the L2ers' overall development because some features might be easier for some L2ers but difficult for others, depending on many factors, such as L1 transfer and exposure to input. This concern is legitimate because it is a

fact that input, L1 transfer, and implicit versus explicit learning (among other factors) can greatly result in a significantly different performance among L2ers (see Section 2.6 below). However, these variables were controlled for in the current study by implementing and following various rigorous criteria when recruiting the L2 participants (see Section 3.2: Chapter 3).

## **2.6 Factors Affecting L2 Development**

In her groundbreaking article (1997), Larsen-Freeman applied the complexity theory, as a purely natural and scientific theory, to SLA. The complex nature of the universe makes it difficult for a single algorithm to provide an in-depth explanation for the various phenomena observed in the universe. Such a theory can thus be extended to language, as a social science, because no single theory can capture the SLA phenomenon (Larsen-Freeman, 1997, 2017). To this effect, over the last three decades, the complexity theory, or the complex dynamic systems theory (CDST) as it has been referred to in the SLA literature, has been relied on as a theoretical framework that could account for the development of L2ers' interlanguage systems (Hiver et al., 2022).

The CDST perceives language as an innately complex process dynamically influenced by a variety of interrelated factors that drive the whole course of interlanguage development (de Bot et al., 2007; Larsen-Freeman, 1997, 2017). One of the main characteristics of the dynamicity of language development is pertinent to the complete interconnectedness of its larger systems (e.g., phonology, lexicon, morphology, syntax), as well as the interconnectedness of the subsystems of each larger system (e.g., the different grammatical structures constituting the whole syntax of a language). However, this does not imply that all connections between (sub)systems are equivalently strong. While the connections for some (sub)systems are tangentially related, the connections for others are quite robust. The development of the interlanguage (sub)systems is contingent on and constrained by limited resources available to L2ers. Those resources can be of

two types: internal and external. The former are those within the L2 individuals (e.g., L1, age, motivation, identity, attitude, gender, aptitude, learning strategies), whereas the latter are those outside the L2ers (e.g., quantity and quality of input, type of instruction, surrounding social environment; de Bot et al., 2007; de Pot & Larsen-Freeman, 2011).

Of all external and internal resources/factors, the current study focuses on input (i.e., external factor), L1 transfer (i.e., internal factor), and implicit versus explicit learning (i.e., internal factors), resulting typically from function-focused instruction and form-focused instruction (i.e., external factors), respectively. The important role of these factors in SLA can be appreciated if they are discussed briefly in relation to FLA. There is general agreement on the fact that FLA is almost always successful, and ultimate attainment of the L1 is reached within the first few years of acquisition. In contrast, SLA does not exhibit this uniformity of success. The three factors mentioned above can offer some accounts for the different outcomes of the two acquisition phenomena.

Regarding the input factor, children typically get exposed to an excessive and modified amount of input in their surrounding environments. On the other hand, adult L2ers usually learn the L2 formally in time-limited settings (e.g., classrooms), which provide them with an insufficient amount of input—potentially leading to less success in the L2 learning process. As for the L1 transfer factor, while children acquire their L1 with zero knowledge of any language, adults commence their L2 learning process with at least one fully developed language system. The L1 system might hinder or facilitate the learning process of L2 features/forms, depending on their presence and absence in the L1, and depending on the L2ers' ability to access the universal grammar (UG: the internally endowed device responsible for acquiring languages) and reconfigure the L1 noninstantiated features onto the L2 (i.e., parameter resetting). The third

factor that differentiates between the two phenomena of FLA and SLA pertains to the learning mechanisms available to young children and adults. FLA takes place implicitly and unintentionally, resulting in knowledge characterized as robust, fluent, effortless, and automatic use of the language. In contrast, SLA is argued to be explicit and conscious, typically resulting in knowledge characterized as slow, effortful, and conscious use of the language. Thorough discussions on the interconnectedness of these factors in SLA are provided further below.

### **2.6.1 Input**

Although some L2 researchers maintain that L2ers resemble L1 acquirers in terms of ending up with mental grammar or internalized language (I-language) that goes beyond the primary linguistic data (i.e., poverty of stimulus; Zyzik, 2009), the majority of SLA researchers emphasize that the quality and quantity of input greatly determine success or failure in SLA (e.g., Alhawary, 2005, 2009a, 2009b, 2013, 2019; Azaz, 2016; Barlow & Kemmer, 2000; Bybee, 2006; Bybee & Hopper, 2001; Corder, 1967; N. Ellis, 2002a, 2002b, 2012; N. Ellis & Robinson, 2008; R. Ellis, 1994; Gass & Mackey, 2002; Gass & Selinker, 2008; Krashen, 1985; Lee, 2002; Long, 1996; Loschky, 1994; Nap-Kolhoff, 2017; Nation & Macalister, 2010; Piske & Young-Scholten, 2009; Sachs et al., 1981; Schmid, 2017; Snow et al., 1976).

According to R. Ellis (1994), SLA scholars, including behaviorists, mentalists, and interactionists, emphasize the need for input in SLA, but they disagree on its role in the acquisition/learning process. Behaviorists perceive language as a set of habits that are developed through stimuli, responses, and reinforcements. According to this perspective, language is acquired by means of imitating (i.e., output) what is heard (i.e., input) and by receiving feedback that can be in the form of reinforcement or correction. Thus, this view rejects any internal processes in the language acquisition device (LAD) and suggests that learning is mainly

constrained and derived by external factors (e.g., input and output) that play the only role in the evolution of the L2 interlanguage system. However, proponents of Chomskyan generative grammar argue for the innate nature of language. Although they do not deny the importance of input for language acquisition, they maintain that its role does not exceed being a “trigger” that activates the internal language mechanisms. The connectionist and interactionist perspective of language takes a neutral position between the aforementioned two schools. It theorizes that both external (e.g., input) and internal factors (e.g., L1) are equally needed in SLA, and it perceives acquisition “as a product of the complex interaction of the linguistic environment and the L2ers’ internal mechanisms, with neither viewed as primary” (R. Ellis, 1994, p. 243).

However, not every input can be beneficial to L2ers. Gass and Selinker (2008) highlighted the important distinction made by Corder (1967) between *input* and *intake*. Input refers to every linguistic data that L2ers are exposed to, including input that “goes in one ear and out of the other” (Gass & Selinker, 2008, p. 305). Intake, on the other hand, refers to data that are comprehended and internalized. Thus, for the input to serve a greater purpose, it must become intake because “it is a given that without understanding the language, no learning can occur. Although understanding alone does not guarantee that learning will take place, it does set the scene for learning to take place” (p. 305).

Nevertheless, because Corder (1967) did not explain what counts as comprehensible input, Krashen (1985) proposed the monitor model, which consists of five hypotheses: the acquisition-learning hypothesis, the input hypothesis, the natural order hypothesis, the monitor hypothesis, and the affective filter hypothesis. According to the input hypothesis, for SLA to take place, L2ers must receive comprehensible input that is slightly above their current level ( $i+1$ ). On this note, Krashen (1985) accentuated the following:



Input is the essential ingredient. The acquirer does not simply acquire what he hears. There is a significant contribution of the internal language processor (Chomsky's LAD). Not all the input the acquirer hears is processed, and the LAD itself generates possible rules according to innate procedures. (pp. 2–3)

Evidence supporting the critical role of input comprehension comes from the FLA literature. For instance, Sachs et al. (1981) and Snow et al. (1976) examined whether receiving input from television only would be beneficial to children in terms of developing comprehension of the language. The findings revealed that those children were unable to internalize the input and transform it into intake—despite the great quantity of input they received through watching television. This strongly denotes that acquisition may not occur without comprehension (as cited in Loschky, 1994, p. 305).

Therefore, some SLA theorists and researchers have suggested many different sources of comprehensible input. For example, Long (1996) emphasized that the interlanguage system and the noticing mechanisms of L2ers can be enhanced and developed through conversational interactions, which entail modifications in the input and output (e.g., recast, simplification). The interactions between proficient users and nonproficient users of a language can facilitate the learning process “because it connects input, internal learner capacities, particularly selective attention, and output in productive ways” (Gass & Mackey, 2002, p. 451).

Likewise, Nation and Macalister (2010) recommended that L2ers can receive comprehensible input through interaction activities that require L2ers to talk with one another, so the output of one L2er can be the input for another and vice versa. Additionally, any L2 course should provide L2ers with a well-balanced range of learning opportunities, so they can receive the best available return for their learning efforts. One way of deciding whether a course provides L2ers with well-balanced opportunities is through perceiving the course as it consists of four strands: meaning-focused input, meaning-focused output, language-focused learning, and

fluency development. The meaning-focused input strand parallels Krashen's input hypothesis. This strand involves learning from listening, speaking, and reading. Three conditions are needed for such learning to take place. First, L2ers must not receive a lot of unknown items at once, even those related to one another in form or function (i.e., opposites, synonyms). Based on several L2 findings, Nation and Macalister (2010) stressed that:

Items which have loose indirect connections with each other (indirect free associates) are learned more effectively if they are learned at the same time. Items which have strong meaning relationships (opposites, near synonyms, free associates) interfere with each other and thus make learning more difficult. (p. 48)

To this effect, teachers are endorsed to adapt the instructional textbook by teaching only one of the linguistic items "that might interfere with each other" (Nation & Macalister, 2010, p. 162). Second, L2ers' attention must be drawn to both the meaning of the message and the form. This dual focus can ensure that L2ers not only understand the messages being conveyed but also pay attention to the syntactic rules underpinning those messages. Third, L2ers must receive a large amount of comprehensible input, which can be achieved by having them read books written specifically for L2ers. Those types of books can ensure higher exposure to comprehensible input and can guarantee that L2ers are not overwhelmed by unknown items, because most of those books introduce a controlled set of vocabulary and forms.

However, input comprehension (i.e., quality of input) *per se* is not enough for new L2 items to be internalized because any input needs to be frequently and sufficiently recycled (i.e., quantity of input). This notion stems from a psychologist perspective postulating that human cognition is generally affected by three experiential factors: frequency, recency, and context. Frequency has a major impact on memory and hence on learning. If people experience something more frequently, their memory become more entrenched about that thing—and thus accessed more easily (N. Ellis, 2002a, 2002b, 2012). To apply this notion to language acquisition

in general, N. Ellis hypothesized that language acquisition is a bottom-up process that is greatly sensitive to input frequency across all language domains, including phonology, reading, spelling, morphosyntax, formulaic expressions, comprehension, production, grammaticality, and syntax. It is the ultimate goal for L2ers to figure out how the L2 works, which can be achieved through frequent exposure to the L2 forms. This is so because “higher frequency words get more activation from the same evidence than do low-frequency words” (N. Ellis, 2002a, p. 151).

The above-mentioned notion is consistent with the entrenchment hypothesis, which predicts that the more lexemes and forms are frequently present in the input, the more likely they are entrenched in the I-language, and hence they are expected to be accessed more easily and used more frequently than others that are not as frequent (Barlow & Kemmer, 2000; Schmid, 2017). Accordingly, from the viewpoint of input frequency proponents, repetition serves as the main mechanism that underpins language representation and acquisition. L2ers are not passive receivers anymore because they can extract regularities and produce systematically linguistic patterns based on the input received (e.g., Barlow & Kemmer, 2000; Bybee, 2006; Bybee & Hopper, 2001; N. Ellis, 2002a, 2002b, 2012; N. Ellis & Robinson, 2008; Schmid, 2017).

The role of input frequency was majorly investigated in the development of FLA and did not receive systematic investigations in SLA until recently. According to Schwartz and Causarano (2007), a potential reason for this lack of attention relates to the fear among SLA theorists and pedagogues of research findings that provide evidence to the input frequency effect. Consequently, this might lead to adopting the behaviorist view of SLA once again and embracing pedagogical approaches that were based on the stimulus-response-enforcement way of instruction (e.g., the audiolingual approach), which promotes memorization of highly frequent constructions. However, “theories of language acquisition are changing, and the wheel’s still in

spin” (N. Ellis, 1998, p. 642), so the important role of input frequency has regained its previous status of importance over the last few decades.

In the FLA (bilingual) literature, input frequency was found to be very influential in the acquisition of an additional language. For example, Nap-Kolhoff (2017) examined the acquisition of the object-naming construction by three Turkish L1 children learning Dutch. Spontaneous data were collected in the children’s homes and their preschool centers. The findings revealed that, despite the overall similarities between the Turkish L1 young children and their monolingual Dutch counterparts in terms of following the same developmental path, some differences were observed. One of the differences was that the Turkish L1 children exhibited a tendency to use only one or two demonstratives in constructions that do not exhibit copula, unlike their Dutch monolingual counterparts, whose average use of demonstratives in these constructions was four demonstrative pronouns. This difference was attributed to many factors, the most important of which was the low quantity of input received by the Turkish L1 children.

In the SLA literature, many studies revealed that the frequent occurrence of the linguistic forms is crucial for acquisition to take place. For example, Lee (2002) examined the effects of frequency on the acquisition of the Spanish L2 future tense by 181 English L1 speakers. The participants read an authentic text under three conditions of exposure: the target form was recycled 6 times in the first condition, 10 times in the second condition, and 16 times in the third condition. The results revealed that those who were exposed to the target form 16 times significantly outperformed those who were exposed to it 10 and 6 times only. Thus, the more L2 forms are repeated in a reading text, the more likely those forms are acquired by L2ers.

Schwartz and Causarano (2007) also explored the effect of gerunds and infinitives frequency on the output of Spanish L2ers of English. To determine the frequency of gerunds and

infinitives in English, the authors used the British National Corpus, which showed that the infinitives were almost 15 times more frequent than gerunds. The results revealed that the participants used the infinitives at a higher accuracy rate, compared to gerunds. Accordingly, the authors argued that the simultaneous exposure to two sets of similar constructions without sufficient recycling should be prevented. Once one of them becomes salient to the L2ers, the other construction could be sequentially introduced. This way of instruction may bring about less ambiguity and less confusion for L2ers.

In the Arabic L2 context, the role of input frequency was likewise reported in many studies. For example, Alhawary (2013) discussed some of the Arabic representative findings, one of which relates to the acquisition of the verb moods assigned by the negation particles: *lā* “do not,” *lam* “did not,” and *lan* “will not.” The data revealed non-emergence or brief emergence of the moods (i.e., indicative, subjunctive, and jussive) due to insufficient recycling of the mood endings in the input (i.e., the instructional textbook). The second representative finding pertains to the acquisition of the nominative case on the subject and the accusative case on the predicates of *kāna* “was” and *laysa* “is not.” The data revealed non-emergence of these case endings in the output of the first-year L2ers due to insufficient recycling of these case-endings in the input (see also Alhawary, 2009a). Based on these findings, Alhawary emphasized that the quality and quantity of input must be taken into account when teaching the case and mood endings to L2ers. This can be achieved by focusing on both forms and functions, as well as by recycling these endings across all language skills—speaking, listening, reading, and writing.

Despite the pivotal role played by input frequency in SLA, it is not the only factor. If the process of learning an L2 revolves merely around acquiring the most frequently recycled forms, then English L2ers, for example, would reach native-like proficiency in their acquisition of the

definite and indefinite articles. According to N. Ellis (2002b), although the definiteness features are considered the most frequently used forms in English, they are considered among the most challenging and difficult features for English L2ers. Therefore, other factors seem to have essential roles to play as well, such as cross-linguistic effects from the L1 (i.e., the initial stage of L2 acquisition). This was well captured by Gass and Mackey (2002), who stated the following:

Research on the role of the native language in SLA has shown that it can impact learning in both positive and negative ways. It is possible that frequency effects are intertwined with native language effects in relation to some aspects of learning. (p. 256; see also Ellis, 2002b; Long & Sato, 1983).

### **2.6.2 L1 transfer**

The existing Arabic L2 studies support the notion that input frequency effects intertwine with L1 transfer effects in the development of Arabic morphosyntactic structures. In addition, it was generally found that “L1 transfer (particularly negative transfer of morphosyntactic features) is noticeably exhibited in lower proficiency levels” more than higher proficiency levels (Alhawary, 2019, p. 138). Accordingly, negative/nonfacilitative L1 transfer seems to be overridden by the proficiency level factor (more exposure to the language). However, before delving into some Arabic L2 representative studies that seem to be in support of these two conclusions, it is crucial to briefly review the most prominent hypotheses that attempt to account for the role of L1 in relation to the notion of UG access.

One of these hypotheses is the full transfer/full access hypothesis (FT/FAH; Schwartz & Sprouse, 1996), which suggests that the L1 grammar is the initial state of L2 acquisition. It predicts that the entire L1 system is transferable from the outset of the learning process, and the UG is fully accessible as well. Thus, native-like competence is attainable, regardless of the structural similarities or differences between the L1 and L2. However, full attainment of native

proficiency is contingent upon transparent exposure to input. If there is obscurity in input, fossilization might be the ultimate outcome.

Another hypothesis is the feature reassembly hypothesis (FRH; Hwang & Lardiere 2013; Lardiere 2008, 2009), which is a reformulation of the missing surface hypothesis (MSH; Lardiere 1998, 2000). It suggests that the learning process goes beyond the simple notion of resetting the parameters of the UG. It assumes that functional features are assembled/bundled in specific and complex ways in different languages. L2ers start learning the L2 while they are equipped with fully assembled L1 features associated with functional categories that most likely differ from those of the L2. The L2 features are ultimately learnable and acquirable, but L2ers must reassemble/reconfigure those functional features from the L1 onto the L2. This reconfiguration is again contingent upon the clarity and transparency of the input received.

A third hypothesis is the interpretability hypothesis (Hawkins 2005; Hawkins et al., 2008), which is a reformulation of the failed functional features hypothesis (FFFH; Hawkins 2001; Hawkins & Chan 1997). It predicts that the L2 acquisition of the uninterpretable  $\phi$ -features (e.g., agreement features) is subject to maturation constraints. That is, acquiring L2 features associated with functional categories that are not instantiated in the L1 is not likely possible after the critical period, regardless of the complex or simple nature of the target features and regardless of the length of exposure to the L2 input.

A very recently proposed hypothesis is the robustness of L1 transfer hypothesis (RL1TH; Alhawary, 2021), which seems to take a mid-position between the hypotheses discussed above. It suggests that the L1 transfer process is selective, depending on the *presence* and *absence* of features in the L1. Based on this prediction, the entirety of the L1 features is not transferable from the outset of the L2 learning process (cf. the FT/FAH). In addition, the RL1TH proposes

that if the L2 features are not instantiated in the L1, they may pose some difficulties for L2ers, regardless of the *easiness* and *complexity* of those features (cf. the FRH and MSH). Moreover, it postulates that the difficulty in the acquisition of the L2 features noninstantiated in the L1, whether *simple* or *complex*, is temporary not permanent (cf. the interpretability hypothesis and FFFH). However, full attainment is contingent upon the frequency and transparency of the input received (à la the FT/FAH and FRH).

The interconnected role of L1 transfer and input quality and quantity was examined in a series of Arabic L2 studies investigating the acquisition of some high-frequency morphosyntactic structures by L2ers. For example, Alhawary (2009a) examined longitudinally the emergence of the subjunctive particle *ʔan* “to” and negation particles in the interlanguage systems of nine English L1 beginner L2ers of Arabic. The findings revealed that the particle *ʔan* “to” emerged in their systems almost 1–2 weeks after it was introduced. Due to the structural proximity between MSA and English in the subjunctive subordination structure, the participants did not encounter any difficulties in supplying the complementizer *ʔan* “to” from early on. As for the negation structures, it was presumed that their mastery depends on two conditions: sequence of presentation (verbal present tense negation, then past tense negation, then future tense negation) and sufficient recycling of these structures in the input. The results revealed that the present tense negation structure with the negator *lā* “does/do not” emerged at an early stage in the interlanguage systems of the participants, who supplied this negator with at least two different verbs. This finding was attributed to the structural proximity between Arabic and English in the present verbal negation and due to its high frequency in the input. As for the past tense negation structure with the negator *mā* “did not,” it was difficult for some participants to acquire, evidenced by the fact that it emerged in their interlanguage systems after the future tense



negation marker *lan* “will not”—although the former was introduced before the latter. The reason for this difficulty was attributed to the insufficient recycling of the particle *mā* “did not” in the participants’ instructional textbook. It was also attributed to cross-linguistic effects because English does not exhibit a negation marker that is similar to *mā* “did not,” which occurs with perfect verbs to convey the past negation. Thus, despite the simple nature of the *mā* + *perfect verb* construction (at least on the surface), it posed some difficulties for the participants due to its *absence* in English, and this result may be argued to support the FT/FAH and RL1TH. Another finding pertains to the acquisition of the other past negator *lam* “did not,” which is exclusively produced with imperfect verbs. This construction did not pose any difficulties for the L2ers due to its *presence* in the participants’ L1 (English). However, despite the sufficient recycling of this negator in the input (i.e., instructional textbook) and despite its structural proximity to its counterpart in English, the participants did not opt for it most often in their production because it was introduced as an optional negator after the particle *mā* “did not” was presented at an earlier stage in the input. Therefore, this finding seems to predictively support the notion that the presentational sequence of certain L2 features in the input plays a crucial role in their acquisition by L2ers. In other words, L2 forms that are presented earlier in the input are argued to be acquired more readily than those introduced at a later stage—even if the later-introduced forms are more similar to their counterparts in the L1 of L2ers and more frequently recycled in the input.

The negation findings reported in Alhawary’s (2009a) study were subsequently confirmed by Albirini et al. (2019), who examined the acquisition of sentential negation (among other structures) by 47 English L1 learners of Arabic as an L2: 25 beginner and 22 advanced participants. Their performance was compared to that of a control group consisting of 17 native

speakers of Arabic. The participants were asked to transform affirmative sentences into their negative counterparts. Overall, the L2 participants (at both proficiency levels) were more accurate in their production of the verbal present negation structure than the past negation structure(s). Precisely, the data of the verbal present tense negation revealed that the mean score of the beginner group was significantly less than that of the advanced group (57.33% and 84.85%, respectively). However, the mean score of the latter was statistically comparable to that of the control participants (100%). As for the past negation structure, the beginner group was significantly less accurate than the advanced group (11.33% and 50.76%, respectively), and both groups were significantly less accurate than the control group (98.04%). The results also revealed that the L2 participants produced the *mā + perfect verb* construction more often than the *lam + imperfect verb* construction—although the latter is more similar to its counterpart in their L1 (English) and more frequently recycled in the input (i.e., instructional textbook). This finding, along with the one reported in Alhawary's (2009a) study, suggests that the role of L1 transfer and input frequency is minimized in the acquisition of the *lam + imperfect verb* construction. More importantly, it provides compelling evidence for the role of the presentation timing of L2 structures in the input. That is, the participants opted for the *mā + perfect verb* construction more often than the *lam + imperfect verb* construction because the former was introduced first, whereas the latter was introduced subsequently as an optional negation structure. This sequence in presenting the two structures might be the main cause that hindered positive/facilitative transfer from the L1.

However, the robust role of L1 transfer and input frequency in Arabic SLA was evidenced by Alhawary's series of studies that examined the acquisition of Arabic Noun–Adjective (N-A) and Subject-Verb (S-V) gender agreement by Arabic L2ers from various L1

backgrounds, including English, French (2005, 2009b), English, French, and Japanese (2009a). The participants were divided based on their L1s (i.e., internal factor) and based on their exposure to Arabic (i.e., external factor): beginner, intermediate, and advanced. The findings showed a positive correlation between accuracy and the internal and external variables in the participants' acquisition of the target structures. Regarding the L1 effects, the French L1 groups were significantly more accurate than the other L1 groups on the N-A gender agreement due to its *presence* in French and *absence* in English and Japanese. As for the S-V gender agreement, all L1 groups performed comparably on the third-person singular masculine form, as well as the third-person singular feminine agreement form, and all L1 groups exhibited a tendency to use the former as the default form in the context of the latter due to the *absence* of S-V gender agreement in their L1s. However, the gender agreement error decreased as the participants progressed in their proficiency (i.e., with more exposure to the Arabic L2 forms). This finding provides evidence to input frequency, which played a critical role in minimizing L1 transfer effects. The sequence of presentation of L2 forms in the input played also an important role in the acquisition of tense and verbal agreement. It was found that, when either one of the past tense or present tense was presented in the input before the other, the one presented first was acquired sooner and produced more accurately than the other tense presented at a later stage. It was likewise found that, when the verbal present agreement was presented before its past counterpart (or vice versa), the participants acquired the one presented first more readily and produced it more accurately than the other tense agreement presented subsequently in the input. These findings thus provide evidence to the FT/FAH—and recently the RL1TH—and counterevidence to the interpretability hypothesis and FFFH because some of the English and Japanese L1 participants eventually acquired the uninterpretable N-A  $\phi$ -feature. That is, some English and

Japanese L1 participants performed at the ceiling level (100% of accuracy), despite the *absence* of the N-A agreement feature in their L1s. Similarly, some English, French, and Japanese L1 participants eventually acquired the uninterpretable S-V  $\phi$ -feature. That is, some participants performed at the ceiling level (100% accurate) on the S-V agreement in both the past and present tense contexts, despite the *absence* of this  $\phi$ -feature in their L1s. These findings were argued to provide counterevidence to the FRH and MSH because the N-A agreement feature was supposed to be easy to learn due to the phonological clue that is realized as a zero morpheme for the singular masculine and as the suffix *-a* for the singular feminine (e.g., *ṭālib jadīd* “a new male student” versus *ṭālib-a jadīd-a* “a new **female** student”). However, the English and Japanese L1 participants had difficulties in acquiring this feature, unlike their French L1 counterparts, who performed at a significantly higher accuracy rate. Clearly, the significant difference between the L1 groups was not due to the *complex* nature of the agreement feature, but it was mainly due to its *presence* in French, on the one hand, and its *absence* in English and Japanese, on the other.

Alhawary (2019) likewise examined the correlated role of input frequency and L1 transfer in the acquisition of N-A and S-V gender agreement, tense/aspect, and null-subject structures by 105 L2ers of Arabic representing two L1 groups: Russian and Chinese. The participants were divided cross-sectionally into four sets of pairings based on their intensive and nonintensive exposure to Arabic: the Chinese L1 group with intensive exposure, Chinese L1 group with nonintensive exposure, Russian L1 group with intensive exposure, and Russian L1 group with nonintensive exposure. Each pairing was further divided into three subgroups based on their proficiency levels: first year, second year, and third year. The N-A gender agreement data revealed that the Chinese and Russian L1 groups with intensive exposure outperformed their counterparts with nonintensive exposure. The outperformance of the intensive groups was

attributed to input frequency effects, given that they received double the amount of input that the nonintensive groups received. Additionally, the Russian L1 participants slightly outperformed their Chinese L1 counterparts on the N-A gender agreement. This finding was attributed to the *presence* and *absence* of the N-A agreement feature in Russian and Chinese, respectively. As for the S-V gender agreement, the Chinese and Russian L1 groups performed comparably in both the past and present tense contexts, despite the fact that Russian exhibits a distinction between third-person singular masculine and feminine in the past tense only. This finding was attributed to the nature of the target structures because both the tense and agreement features are conflated in the prefix for the present tense and in the suffix for the past tense—and the L1 of these participants exhibit a tense distinction between past and present. Therefore, a conclusion could not be reached regarding the role of L1 transfer in the acquisition of S-V gender agreement due to the difficulty in ascertaining whether these participants actually acquired the agreement feature (in addition to the tense feature) or they acquired the tense feature only. Moreover, the poor performance of the Chinese L1 groups on the N-A gender agreement compared to their good performance on the S-V gender agreement was argued to provide counterevidence to the input frequency effect. That is, despite the fact that Chinese exhibits neither N-A agreement nor S-V agreement, the Chinese L1 groups did not have problems with the latter, as they did with the former. If input frequency was at stake here, then they would have performed comparably on both structures, which are highly frequent in the input. Instead, this asymmetrical performance on both agreement features highlights again the role of L1 transfer. In other words, the *absence* of the N-A gender agreement in Chinese was most likely the cause for the difficulty encountered by the Chinese L1 groups, and the *presence* of the tense feature in Chinese and Russian perhaps assisted the participants from both L1 backgrounds to produce the verbs in a target-like way.

The findings from the tense data may provide evidence to the important role of presentation timing of the input more than input frequency and L1 transfer. Because Chinese and Russian exhibit a tense distinction between past and present, positive transfer and comparable performance were expected from the Chinese and Russian L1 groups from early on (as in Al-Hamad's (2003) study, which found Arabic L2ers from these two L1 backgrounds to be very successful in their production of both tense structures). However, the Chinese first-year and second-year participants with nonintensive exposure underperformed their Russian counterparts on the past tense because it was formally introduced in the textbook used by the Chinese L1 participants at a later stage compared to the textbook used by the Russian L1 participants. Similarly, the Russian L1 first-year and second-year participants with nonintensive exposure and the Russian L1 first-year participants with intensive exposure underperformed their Chinese L1 counterparts on the present tense because it was introduced at a later stage in the former participants' textbook. Thus, the underperformance of the Chinese and Russian L1 participants on the past tense and present tense, respectively, and the outperformance of the Chinese and Russian L1 participants on the present tense and past tense, respectively, were most likely due to the presentation timing of both tense structures in the input (i.e., instructional textbooks). A similar effect of input exposure was yielded from the tense data in Alhawary's (2009a) study, which found that English, Spanish, and Japanese L1 learners of Arabic as an L2 were more accurate in their production of the past tense than the present tense because the former was presented in the input much earlier than the latter.

The role of L1 transfer was strongly supported by the null-subject data in Alhawary's (2019) study. The Chinese third-year participants (with intensive and nonintensive exposure) dropped subjects at a higher rate than their Russian counterparts, who were more conservative in

dropping subjects. Given that Chinese is like Arabic in terms of being a null-subject language and that Russian is a mixed null-subject language, the difference in performance between the two L1 groups was attributed to L1 transfer effects rather than input frequency effects. In other words, the *presence* of the null-subject feature in the Chinese participants' L1 assisted them in approximating further toward the target language. In contrast, the mixed null-subject feature in Russian did not facilitate dropping subjects more freely by the participants from this particular L1 background, regardless of the prolonged and intensive exposure to Arabic. Similarly, the data of the N-A gender agreement feature revealed an advantage for the Russian L1 participants over their Chinese L1 counterparts due to the *presence* and *absence* of this uninterpretable  $\phi$ -feature in the L1s of the former and latter groups, respectively. Thus, contrary to the interpretability hypothesis, FRH, and MSH, which seem to fall short in explaining the data, the findings can be best accounted for by the FT/FAH and RL1TH because some Russian and Chinese participants performed at ceiling on the null-subject and N-A agreement features, respectively—despite the *absence* of the former in Russian and the latter in Chinese.

The FRH is supported by the findings reported by Azaz (2016), who examined the acquisition of the Arabic definiteness features by English L1 beginner and advanced Arabic L2ers. Before collecting the data, the author hypothesized the following: (a) the participants would not encounter difficulties with Determiner Phrases (DP) that denote unique entities (e.g., *ʔaṭ-ṭaqs* “the weather”) because these DPs are definite in English; (b) the participants would encounter difficulties with DPs that denote abstract nouns (e.g., *ʔaṣ-ṣadāqah* “friendship”) because English abstract nouns are bare; (c) the participants would fluctuate between definite and indefinite use in DPs that denote generic nouns (e.g., *ʔal-kalb* “a dog/the dog”) because English generic nouns can be definite or indefinite. Data were collected from a grammaticality

judgment task, where the participants judged the acceptability of 35 sentences by identifying and correcting the errors. Data were also gathered through a forced choice task, where the participants chose whether the nouns in 40 sentences were definite or bare nouns. The findings revealed that L1 influence was afoot in processing the Arabic definiteness feature because the performance of the beginner participants on unique entities, in which English and Arabic converge, was better than their performance on abstract nouns, in which the two languages diverge. The findings also revealed that the advanced participants exhibited a steady and target-like performance on Arabic generic nouns. Thus, the interpretability hypothesis and FFFH were falsified because the participants were able to reconfigure the DP definite feature denoting generic nouns from the L1 (English) onto the L2 (Arabic) as their proficiency increased, which in turn provides evidence to the FRH. The inability of the beginner participants to reassemble this complex feature was likely due to the lack of explicit instruction, taking into consideration that they used *Al-Kitāb fī Ta'allum Al-'Arabiyya* (Brustad et al., 2011), which does not explicitly point out the similarities and differences between the two languages in terms of the use of the definite article. Thus, this finding highlights the important role of explicit instruction in the acquisition of L2 features, particularly those that are absent in the L1 or set/configured asymmetrically in the L1 and L2.

### **2.6.3 Implicit and explicit learning**

Input frequency and L1 transfer are not the only significant contributing factors in SLA, as the notion of implicit versus explicit learning seems to play a central role as well. According to R. Ellis (2009a), implicit learning “takes place without either intentionality or awareness,” whereas explicit learning “is necessarily a conscious process and is generally intentional as well” (p. 7; see also DeKeyser, 2003; N. Ellis, 2005, 2015; R. Ellis, 1994).



Implicit learning generally results from implicit instruction that propels L2ers to infer the L2 rules incidentally (i.e., without awareness and without attending to the L2 forms) through spontaneous meaning-oriented activities and through presenting the L2 forms in context. The ultimate outcome of this type of learning is most often implicit knowledge *of* the language itself, which could bring about automatic, effortless, fluent, subconscious, and intuitive use of it. In contrast, explicit learning typically results from instruction that focuses on presenting the L2 rules deliberately (i.e., with consciousness and awareness of L2 forms) by means of using meta-linguistic terms and through controlled form-oriented activities. This type of learning usually leads to explicit knowledge *about* the language (i.e., meta-linguistics), which may debatably result in nonautomatic, effortful, slow, conscious, and aware use of it (e.g., DeKeyser, 2003; N. Ellis, 2005, 2015; R. Ellis, 1994, 2009a; Long & Robinson, 1998; Norris & Ortega, 2001).

However, although implicit learning produces more robust (implicit) knowledge, the mechanisms of such learning are argued to be very limited among adult L2ers due to cognitive maturation (i.e., learning after the critical period), which leads them to compensate for this deficit by learning the L2 explicitly (DeKeyser, 2003; DeKeyser & Larson-Hall, 2005). Even if one rejects this claim, implicit learning takes longer than its explicit counterpart, as in FLA (DeKeyser, 2003; Ellis, 2002a). To this end, L2 instruction often tends to be deliberate and explicit because it helps L2ers learn quickly in time-limited situations (e.g., classroom settings). Although it is the ultimate goal to develop implicit L2 knowledge, an important question remains as to whether explicit learning can be eventually converted into implicit knowledge. There are at least two positions regarding this issue: the noninterface position and the interface position (e.g., DeKeyser, 2003; N. Ellis, 2005; R. Ellis, 1994, 2009a; Godfroid, 2022).

The noninterface position stems from Krashen's (1985) monitor model, where he made a distinction between the terms *acquisition* and *learning*. On the one hand, acquisition is typically referred to as a subconscious process indistinguishable from the process of FLA, in which the linguistic system develops implicitly and procedurally—resulting in effortless and spontaneous language production. On the other hand, learning is a conscious process that leads to meta-linguistic explicit knowledge about language—potentially generating effortful and slow utterances. Based on this distinction, Krashen (1985) is an advocate of the noninterface position, which postulates that the conversion of explicit learning into implicit knowledge is not likely possible. According to this view, intentionally learned rules function as a *monitor* that hinders the automaticity and speed of normal communication.

In contrast, the interface position, informed by the noticing hypothesis (Schmidt, 1990), argues that the conversion of *input* into *intake* is contingent upon noticing it consciously. This position predicts that explicitly learned rules can turn into implicit knowledge through controlled practice, form-focused instruction (Norris & Ortega, 2001; Robison, 1996; Spada & Tomita, 2010), and visual presentation of input (Al Midhwah & Alhawary, 2020; Kim & Godfroid, 2019; Lin & Alhawary, 2018). In fact, these studies have found explicit learning to be even more effective than implicit learning in the development of implicit knowledge among L2ers in time-limited classroom settings.

Robison (1996) examined the role of implicit and explicit learning in English L2ers' grammaticality judgment speed and accuracy on the complex rule of pseudo-clefts of location and the simple rule of S-V inversion after adverbials. The results revealed that the implicit-mode L2ers and explicit-mode L2ers performed comparably on the complex rule, but the explicit-mode L2ers significantly outperformed their implicit-mode counterparts on the simple rule.

The advantage of explicit learning over its implicit counterpart in producing more robust (and perhaps implicit) knowledge in L2 settings was confirmed by Norris and Ortega (2001), who provided a meta-analytical and systematic analysis—based on the effect size scale—of 77 experimental L2 studies that examined the efficacy of different types of L2 instruction, two of which were the implicit and explicit types of instruction. Overall, the analysis revealed that the instructional incorporation of explicit techniques was more effective than implicit instruction because the former corresponded to significantly larger effect sizes than the latter (for a similar analysis and result, see also Spada & Tomita, 2010).

The advantage of repeated *visual* presentation of input (i.e., explicit input enhancer) over its *aural* counterpart (i.e., implicit input enhancer) was recently reinforced by Kim and Godfroid (2019), who examined the efficacy of the visual and aural modalities in developing implicit syntactic knowledge among L2ers. One of the study experiments was conducted on 61 English L1 speakers. It consisted of three sets: training, testing, and debriefing. The stimulus was in the form of an artificial language that exhibited the V-S word order from German and the lexicon from English. For the training set, 120 sentences were presented *aurally* to 30 participants, and the same sentences were presented *visually* to the remaining 31 participants. For the test set, 60 sentences were generated. Both groups (visual and aural) were instructed that the 120 sentences in the training set followed complex rules and were asked to listen to or read the new sentences to identify whether they conformed to the complex rules. Subsequently, they were asked to indicate the source of their judgment based on the following criteria: guess, intuition, recollection, or rule. If their performance was above chance (i.e., 50% or above) and was based on guess or intuition, then their knowledge was considered implicit. If their performance was above chance and was based on recollection or rule, then their knowledge was considered

explicit. Upon completion of the test phase, the participants completed a debriefing questionnaire asking them whether they noticed any regularities. If their responses were positive, they were asked to indicate from which set and mode they extracted the regularities. The overall results revealed that the aural mode (i.e., focus on function) was less effective than the visual mode (i.e., focus on form). In addition, although both groups of participants developed explicit knowledge, only the visual-mode group developed a more robust implicit type of knowledge.

The role of implicit and explicit learning in the Arabic L2 literature has not received enough attention (Alhawary, 2018), but there have been several studies highlighting the importance of explicit learning in developing implicit knowledge and emphasizing the critical role of typographically visual input, especially with Arabic L1 features that are absent in the L1 or need to be reconfigured from the L1 onto the L2. One of these studies was carried out by Lin and Alhawary (2018), who did not directly examine the role of explicit learning in the development of implicit knowledge, as their focus was on the role of input frequency and L1 transfer in comprehending and producing Arabic lexical stress by English L1 and Chinese L1 learners of Arabic as an L2. However, the results can contribute to the literature on explicit and implicit learning. The findings found frequency effects on producing Arabic lexical stress by these L2ers in the production task. The participants were able to produce stress more accurately and fluently with trisyllabic words (i.e., 'CvCvCv), which are more frequent in the input than other syllabic patterns. However, the L1s of the L2ers played a more important role than input frequency in their perception and production of stress. This was evident in the significant outperformance of the Chinese L1 participants over their English L1 counterparts because the former participants were positively influenced by Chinese, which heavily relies on stress and tone. However, the performance of the Chinese L1 participants was significantly different from

that of the control group—despite the presence of stress in Chinese. The indication of this result is that implicit learning (aural mode) of Arabic stress did not turn into implicit knowledge. Therefore, the authors suggested raising the L2ers' conscious awareness of Arabic stress through integrating form-oriented drills that explicitly focus on developing the L2ers' production and comprehension of stress, especially in the absence of visual presentation of stress in Arabic. Those controlled activities not only make the L2ers attend to the stress feature but also ensure sufficient recycling of it in the input.

The advantage of explicit learning over implicit learning was likewise confirmed by Al Midhwah and Alhawary (2020), who examined the role of diacritics in word recognition, reading speed, accuracy, and comprehension among 54 English L1 beginner, intermediate, and advanced L2ers of Arabic. They were equally divided into two main groups: the first consisted of L2ers exposed to vowelized materials, and the second consisted of L2ers exposed to unvowelized materials. The participants completed three tasks (word-list reading, text reading, and target-word comprehension) under vowelized and unvowelized conditions. The findings revealed that the vowelized groups significantly outperformed their unvowelized counterparts in their reading speed, accuracy, and comprehension across all proficiency levels and in both the vowelized and unvowelized conditions. The outperformance of the vowelized groups was attributed to the repeated presence of internal vowels (i.e., diacritics) in their textbook, which provided them with “a more advantageous (visual) input” (p. 435), compared to their counterparts with exposure to unvowelized materials. In addition, there was a conscious in-class attempt by the teachers of the vowelized groups to draw their students' attention to diacritics. Hence, the findings provide evidence for the superiority of visual input (explicit) over aural input (implicit) in developing implicit knowledge—thereby supporting the interface position.

## **2.7 Summary and Implications of the Three SLA Factors for the Current Study**

The diglossic status of Arabic (Ferguson, 1959a) or the existence of a continuum of Arabic varieties (Badawi, 1973; Blank, 1960; Mahmoud, 1986; Meiseles, 1980; Mejdell, 2021; Mitchell, 1986; Ryding, 1991) have propelled Arabic L2 programs to embrace various pedagogical approaches that aim to equip Arabic L2ers with linguistic and sociolinguistic competencies similar to those of native speakers of Arabic. The most widely adopted approach is to teach MSA only for the first 2–3 years, followed by studying in courses dedicated solely to the learning of a CA, as well as opportunities for L2ers to study abroad in Arabic-speaking countries, where they can receive an adequate amount of authentic input in CA (Alhawary, 2013, 2021; Eisele, 2018; Ryding, 2006).

However, the MSA-then-CA approach has been challenged because of the assumption that it does not reflect the sociolinguistic reality of the Arab world, in which MSA and CA coexist side by side in complementary distribution (Al-Batal, 1992, 1995, 2018; Younes, 1990, 2015), and because it creates a “firewall separation vision” between MSA and CA (Al-Batal, 2018). To this end, Younes (1990, 2015) proposed the so-called IA, which calls for the simultaneous teaching and learning of the two varieties from early on (see also Al-Batal, 1992, 1995, 2018). This approach has been momentarily supported by others, who found—based on questionnaire surveys—positive attitudes held by Arabic L2ers toward the simultaneous learning of MSA and CA (Al-Batal & Glakes, 2018; Belanp, 2006; Hashem-Aramouni, 2011; Huntley, 2018; Husseinali, 2006; Isleem, 2018; Kuntz & Benlap, 2001; Palmer, 2007, 2008; Shiri, 2013).

However, opponents of the IA maintained that such an approach does not address the four elements constituting the core of any pedagogical approach: understanding the nature of language, how an L2 is learned, the different methods and techniques that promote language learning, and how to assess language learning outcomes (Alhawary, 2021). Additionally, the IA

seems to neglect the four noncontroversial principles that derive the course of SLA: clarity of input, frequency of input, integration of form and function, and integration of all language skills (Alhawary, 2013). Moreover, the IA does not provide any practical solutions to the linguistic dissonance experienced by L2ers who are instructed in two varieties from the outset of their learning process (as reported in Al-Batal & Glakes, 2018; Huntley, 2018). If confusion is experienced by L2ers, they most likely do not comprehend the L2 input, which is considered a prerequisite for language acquisition to take place (Alhawary, 2013, 2021; Corder, 1967; Gass & Mackey, 2002; Gass & Selinker, 2008; Krashen, 1985; Loschky, 1994; Long, 1996; Nation & Macalister, 2010; Sachs et al., 1981; Snow et al., 1976). Additionally, it was found in many studies that input incomprehensibility is usually the outcome of receiving optional yet distinct forms of the L2, which most likely results in problematic production (as in Albirini et al., 2019; Alhawary, 2013, 2009a; Schwartz & Causarano, 2007).

Only three studies examined the learning outcomes of IA. Leddy-Cecere (2018) and Nassif (2018) found L2ers who were exposed to MSA and CA simultaneously developed native-like linguistic and diglossic competencies (i.e., code-switching) as their proficiency increased. Similarly, Shiri and Joukhadar (2018) examined the code-switching behavior among beginners learning MSA and CA simultaneously but in different class sections. It was found that such a pedagogical approach assisted the L2ers in attaining knowledge of the differences between MSA and CA and recaching a high level of accuracy in their production. However, these studies came with many methodological limitations, the most important of which is that they relied heavily on presentations and dialogue skits to collect the data. Such instruments lack spontaneity, which is considered one of the most important indicators of competence development (Alhawary, 2009a, 2019, 2021) and implicit knowledge (R. Ellis, 2009b).

Given the paucity of studies examining the simultaneous learning of MSA and CA from the outset, as well as the methodological flaws of the available studies, it is crucial to fill in this gap in the literature by comparing the spontaneous output of Arabic L2ers who are exposed to two varieties versus those who are exposed to one variety. As suggested by Alhawary (2018) and Wahba (2023), such a comparison can, in turn, provide empirical evidence to determine which one of the two pedagogical approaches (i.e., teaching MSA-then-CA versus teaching MSA and CA simultaneously) yield superior learning gains.

As predicted by the CDST, development of the L2 (sub)systems is restricted by the limited resources available to L2ers (de Bot et al., 2007; de Pot & Larsen-Freeman, 2011; Larsen-Freeman, 1997, 2017). Some of these resources (or factors) interconnect strongly with one another, particularly input, L1 transfer, and explicit learning. Input comprehension and frequency are essential for SLA to take place (e.g., Albirini et al., 2019; Alhawary, 2005, 2009a, 2009b, 2013, 2019; Azaz, 2016; Barlow & Kemmer, 2000; Bybee, 2006; Bybee & Hopper, 2001; N. Ellis, 2002a, 2002, 2012; R. Ellis, 1994; Gass & Mackey, 2002; Gass & Selinker, 2008; Krashen, 1985; Lee, 2002; Long, 1996; Loschky, 1994; Nation & Macalister, 2010; Piske & Young-Scholten, 2009; Sachs et al., 1981; Schmid, 2017; Snow et al., 1976). However, it does not work in isolation from other essential factors, such as L1 transfer (Alhawary, 2005, 2009a, 2009b, 2013, 2019; Azaz, 2016; N. Ellis, 2002b; Gass & Mackey, 2002; Lin & Alhawary, 2018; Long & Sato, 1983). The notion of implicit (incidental) and explicit (intentional) learning seems to play a core role as well (DeKeyser, 2003; N. Ellis, 2005, 2015; R. Ellis, 1994, 2009a; Long & Robinson, 1998; Norris & Ortega, 2001). In L2 time-limited classroom settings, explicit learning seems to be more effective than its implicit counterpart in developing implicit knowledge (i.e., robust, fluent, effortless, and automatic language use). Specifically, reliance on deliberately



form-focused (explicit) instruction has an advantage over spontaneously function-focused (implicit) instruction (Norris & Ortega, 2001; Robison, 1996; Spada & Tomita, 2010). In addition, visual input, as an explicit input enhancer, is more advantageous to L2ers than aural input, as an implicit input enhancer (Al Midhwah & Alhawary, 2020; Kim & Godfroid, 2019; Lin & Alhawary, 2018).

In light of the SLA factors reviewed in the preceding sections, three broad possibilities are obtained with regard to the simultaneous teaching and learning of two Arabic varieties from early on. First, input in both varieties (MSA and EA) must be comprehended and sufficiently recycled, especially with forms that are absent in the L1 of the L2ers or need to be reassembled or reconfigured from the L1 onto the L2. Without repeated exposure to input in all language skills, learning might not take place, and L2ers would most likely produce hybrid constructions that are not target-like (as in Leddy-Cecere, 2018). However, it is not likely possible to repeatedly and frequently expose Arabic L2ers to all MSA and EA input in those courses due to the limited time of Arabic courses offered in U.S. universities (2–3 years), during which Arabic L2ers can barely develop “basic competence (Abboud, 1968) and “working knowledge” (Alhawary, 2013, 2021) of the morphosyntactic structures of one Arabic variety only. Furthermore, optionality in the input without sufficient recycling might prevent comprehension, which can result in output that is nontarget-like. Thus, different performance is expected to be the outcome between Arabic L2ers who are instructed in MSA and EA simultaneously versus those who are instructed in MSA only.

Second, Arabic L2ers might not encounter any difficulties in processing and eventually producing the MSA and EA target structures (see Section 3.3, Chapter 3), but this is contingent on the presence of the target features in their L1 (English), as well as on their ability to

reconfigure the noninstantiated features from their L1 onto the L2. Additionally, the reconfiguration and full attainment of the absent or noninstantiated features is contingent on the transparency of input, as predicted by various L1 transfer hypotheses, particularly the FT/FAH (Schwartz & Sprouse, 1996), FRH (Hwang & Lardiere, 2013; Lardiere, 2008, 2009), MSH (Lardiere, 1998, 2000), and RL1TH (Alhawary, 2021). Given that exposure to two Arabic varieties violates the clarity of input principle (Alhawary, 2013; see also Nation & Macalister, 2010; Schwartz & Causarano, 2007), L2ers who receive input in MSA and EA simultaneously from early on might experience confusion, which might hinder positive/facilitative transfer from their L1 (English) and make their access to UG relatively more difficult than L2ers who receive transparent input in one variety only. Thus, different performance between the two groups of Arabic L2ers is perhaps the outcome.

Third, due to the effectiveness of explicit learning of L2 grammatical structures and its positive impact on raising L2ers' conscious awareness about the gaps in their interlanguage systems, form-oriented drills in both varieties (MSA and EA) should be sufficiently incorporated into the Arabic curricula. However, given that all CA varieties (including EA) are not fully codified (Wahba, 2023), this entails negligence of explicitly form-focused instruction of EA morphosyntactic forms (Alhawary, 2013)—potentially resulting in a difficulty and delay in internalizing those forms. In addition, given the vital role of visual presentation of input (as an explicit technique), MSA and EA input should be introduced to Arabic L2ers visually, not just aurally. However, because EA is not used in writing, L2ers might have very limited exposure to visual input in this variety. As a consequence, they would be left with aural input only, which might be less effective in acquiring the EA structures (as in Al Midhwah & Alhawary, 2020; Kim & Godfroid, 2019; Lin & Alhawary, 2018).

## Chapter 3 Methodology

### 3.1 Overview

The current study aimed to examine the efficacy of integrating two varieties in teaching and learning Arabic as an L2. It did so by measuring and comparing the learning outcomes (in terms of fluency, accuracy, complexity, and overall proficiency) of two groups of 57 English L1 learners of Arabic as an L2: those who were instructed in Modern Standard Arabic (MSA) and Egyptian Arabic (EA) simultaneously from the outset, the integrated program group (IPG), and those who were instructed in MSA only, the unintegrated program group (UIPG), across three proficiency levels: first year, second year, and third year. The performance of both L2 groups was compared to that of a control group (CG), consisting of nine native speakers of EA.

The two research questions (RQs) proposed in Chapter 1 are restated below:

**RQ1.** What is the nature of the language output of Arabic L2ers in the IPG versus that of Arabic L2ers in the UIPG across the three proficiency levels (first year, second year, and third year)? More specifically:

- Are there any similarities and differences between the IPG and UIPG in terms of their fluency when narrating past and present events?
- Are there any similarities and differences between the IPG and UIPG in terms of their accuracy on the target structures of past tense and present tense, as well as past negation and verbal present negation?

- Are there any similarities and differences between the IPG and UIPG in terms of their error-free production of the target complex structures of subjunctive subordination and cause–effect subordination?
- Are there any similarities and differences between the IPG and UIPG in terms of their overall proficiency?

**RQ2.** What is the role of input, L1 transfer, and implicit versus explicit learning in affecting the performance of the IPG versus that of the UIPG?

To answer these questions, spontaneous obligatory-in-context data were collected from all participants in elicitation sessions that were held in the form of in-person one-on-one interviews. The participants performed three oral picture-based description tasks: two narrative tasks and one manipulation task. The nature of the current study entailed analyzing the data quantitatively and qualitatively. Quantitative comparisons and analyses were conducted to detect whether there were any statistically significant effects of the type of participant groups and the proficiency levels (i.e., independent variables) on the participants' fluency, accuracy, complexity, and overall proficiency (i.e., dependent variables). Moreover, the correct and incorrect rule application by all participants when producing the target structures was qualitatively analyzed to lay the groundwork for examining the role of input, L1 transfer, and implicit versus explicit learning in the interlanguage development of the target structures among the IPG and UIPG participants across the three proficiency levels of first year, second year, and third year.

The current chapter provides detailed discussions of the participants, target structures in MSA and EA, materials, data collection procedure, coding criteria, and data analysis.

## **3.2 Participants**

A high level of compatibility and homogeneity between the two L2 groups at each proficiency level was vital to reach valid results and solid evidence. Accordingly, a number of relevant variables had to be tightly controlled for through strict adherence to certain criteria developed for the Arabic L2 program selection and participant recruitment. Sections 3.2.1, 3.2.2, and 3.2.3 provide detailed discussions of the criteria considered in selecting the Arabic L2 programs, along with information about the two selected programs, the recruitment criteria and results of the L2 participants, and the recruitment criteria and results of the control participants.

### **3.2.1 Arabic L2 program selection**

Two Arabic L2 programs were selected as the two research/recruitment sites: integrated program (IP) and unintegrated program (UIP). In selecting these two requirement sites, many compatibility measures were taken into account to control for several variables related to the general quality of education of the two universities to which the two Arabic programs belonged, program structure, medium of instruction, and instructional textbooks.

#### ***3.2.1.1 University rankings***

To control for the quality of academic education that the L2 participants in both programs received, the two Arabic programs had to belong to U.S. universities whose academic rankings were similar in the academic year during which the data collection took place. Selection of the two schools relied on the U.S. News Report for Best Colleges Rankings (2021) because it implemented various indicators that reflect the academic quality of faculty members and students. Those indicators include graduation rate performance, social mobility, faculty resources, expert opinion, financial resources, student excellence, and alumni giving.

Accordingly, the two selected Arabic programs were offered at two prestigious U.S. universities,

whose rankings were within the first 70 institutions across the states during the time of data collection (U.S. & World Report, 2021).

### ***3.2.1.2 Program structure***

The two selected programs were similar in enrollment size and instruction length. Both had Arabic L2ers who met the L2 participant recruitment criteria (see Section 3.2.2 below). Both focused on developing all language skills—speaking, listening, reading, and writing—across three proficiency levels: first year, second year, and third year. Each level was divided over two semesters, and the length of each was 16 weeks. In each semester, the IP offered four credits per week for each proficiency level. Similarly, the UIP offered four credits per week for the first-year and second-year levels and only three credits for the third-year level.

### ***3.2.1.3 Medium of instruction***

Instructionally, the IP adopted the two varieties of MSA and EA as two mediums of instruction across the three proficiency levels. Thus, all content and drills were presented to the IPG in both varieties from the onset of their learning and maintained throughout. In contrast, the UIP adopted the MSA variety as the only medium of instruction across the three proficiency levels. Thus, all contents and drills were presented to the UIPG in MSA only and maintained throughout.

### ***3.2.1.4 Instructional textbooks***

For the curriculum, the IP and UIP used the textbook of *Alif Bā'*: *Introduction to Arabic Letters and Sounds* (Brustad et al., 2010) during the first half of the first semester in first-year Arabic classes. This textbook provides Arabic L2ers with the basic materials needed to learn the Arabic alphabet, sounds, and vocabulary. Both programs also used the same series of *Al-Kitāb fī Ta'allum Al-'Arabiyya* (Brustad et al., 2004, 2007, 2011, 2013). However, they adopted two

different editions, each of which has a distinct presentation of units/lessons and asymmetric focus on the Arabic varieties. Table 3.1 displays the series edition used by each program and the units covered at each proficiency level.

The IP used Part I and Part II of the third edition of *Al-Kitāb fī Ta'allum Al-'Arabiyya* (Brustad et al., 2011, 2013) at all levels of Arabic classes. What distinguishes this series from other Arabic textbooks is its integration of MSA, EA, and Levantine Arabic (LA) from the beginning and throughout the series. It consists of 23 units/lessons (13 in Part I and 10 in Part II). Besides the textbook of *Alif Bā'* (Brustad et al., 2010), 10 units of Part I of this edition were covered by the end of first-year classes; the remaining units of Part I and 5 units of Part II of this edition were covered by the end of second-year classes; the remaining units of Part II was covered by the end of the first semester of third-year classes. Each unit is structured as follows. First, the MSA, EA, and LA vocabularies are concurrently presented. Next, a story is presented in either EA or LA through a video. The same story is then repeated in MSA. From there, the instruction of grammar focuses mainly on MSA, occasionally followed by a brief explanation of the grammatical differences between MSA, EA, and LA. Reading materials are presented subsequently in MSA only. Lastly, an EA or LA dialogue is presented aurally and followed by some drills to enhance speaking.

The UIP used Part I and Part II of the second edition of *Al-Kitāb fī Ta'allum Al-'Arabiyya* (Brustad et al., 2004, 2007) in the second-year and third-year Arabic classes. This edition focuses on the explicit instruction of all language skills and components in MSA. Although some concepts and materials from CA (mainly EA) are incorporated in this edition, those materials are presented briefly in the last two pages of each unit/lesson. This edition consists of 30 units (20 in Part I and 10 in Part II). The last 7 units of Part I and three units of

Part II of this edition were covered by the end of second-year classes; the remaining units of Part II were covered by the end of the first semester of third-year classes. As for the first-year classes, the third edition of the *Al-Kitāb* series was used instead of the second edition. Besides the textbook of *Alif Bā'* (Brustad et al., 2010), 11 units of Part I of the third edition were covered by the end of first-year classes.

**Table 3.1** *The Edition and Parts of Al-Kitāb Used in the IP and UIP, as well as the Number of Units Covered at Each Proficiency Level*

Group and Proficiency Level	Weekly Credits/Hours	Al-Kitāb Edition/Part	Lessons Covered
<b>IP</b>			
1st year	4	Third Edition/Part I	1–10
2nd year	4	Third Edition/Parts I & II	11–13 (I) & 1–3 (II)
3rd year	4	Third Edition/Part II	4–10
<b>UIP</b>			
1st year	4	Third Edition/Part I	1–11
2nd year	4	Second Edition/Parts I & II	14–20 (I) & 1–3 (II)
3rd year	3	Second Edition/Part II	4–10

*Note*<sup>1</sup>. IP = integrated program; UIP = unintegrated program.

*Note*<sup>2</sup>. Both programs used the textbook of *Alif Bā'*: *Introduction to Arabic Letters and Sounds* (Brustad et al., 2010) during the first half of the first semester in the first-year Arabic classes.

### 3.2.2 L2 participants

To obtain general background information about the L2 participants, they were administered the Language History Questionnaire (LHQ; see Appendix A). It involved questions about their age, gender, L1, education level, knowledge of L2s other than Arabic, prior exposure to Arabic through travels to Arabic-speaking countries, prior exposure to Arabic through previous enrollments in other programs, exposure to CA, frequency of code-switching behavior between MSA and CA, year level in the current Arabic program, and self-rating of Arabic proficiency.

Among all information collected from the LHQ that indicated some variabilities among Arabic L2ers, only the following variables were taken into account for the L2 participant



inclusion and exclusion criteria: L1 background, prior exposure to Arabic, exposure to CA, and proficiency level. The recruitment criteria and results are thoroughly discussed in Sections 3.2.2.1 and 3.2.2.2, respectively.

### ***3.2.2.1 L2 participant recruitment criteria***

#### *3.2.2.1.1 L1 background of the L2 participants*

The first important criterion considered in the recruitment process of the L2 participants was their L1. As will be discussed in Section 3.3 below, MSA, EA, and English converge and diverge in the target morphosyntactic structures examined in this study. To examine the role of L1 in the acquisition of the target structures in MSA and EA, any cross-linguistic effects from languages other than English had to be controlled for.

Accordingly, two types of Arabic L2ers in the two L2 research sites were excluded from the recruitment: those who reported any language other than English as their L1 and those who were Arabic heritage speakers (i.e., L2ers of Arabic ancestors). The latter were excluded for two reasons. First, they highly likely had prior exposure to Arabic through their parents and siblings, which most likely provided them with an advantage over other nonheritage L2ers. Second, CA was considered their L1, along with English (Albirini, 2019), so their proficiency in CA can influence their proficiency in MSA and vice versa, as empirically evidenced by Albirini (2018).

#### *3.2.2.1.2 Prior exposure to Arabic*

The second essential criterion considered in the recruitment process of L2 participants pertained to their prior exposure to Arabic. The L2 participants in each subgroup of the three proficiency levels—first year, second year, and third year—had to be homogenous in their respective proficiency levels. That is, the L2 participants representing each year in both programs were supposed to receive a comparable amount of Arabic input. To control for the variable of prior

input exposure, the LHQ included the following two questions: (a) Have you ever learned Arabic before joining your current program? and (b) Have you ever lived in an Arabic-speaking country? Those who provided a positive answer to the two questions were further asked to report the duration of their prior learning and the duration of their stay in Arabic-speaking countries. The choices provided in the LHQ were as follows: 1–3 months, 4–6 months, and more than 6 months.

Accordingly, two types of Arabic L2ers in the two L2 recruitment sites were excluded: those who had previously studied in other Arabic language programs for more than 6 months and those who had traveled to Arabic-speaking countries and resided there for more than 3 months. These two types of L2ers were excluded because their prior exposure to Arabic input (either formally or informally) would have likely provided them with an input advantage over their L2 peers who had not had the same prior exposure. It would have been ideal to recruit L2 participants who had never been exposed to Arabic before joining their current programs. However, this was not likely possible due to the relatively limited number of participants who agreed to take part in the study in both recruitment sites. Therefore, those who had studied Arabic in previous programs for a duration up to 6 months and those who had lived in Arabic-speaking countries for a duration up to 3 months were recruited to come up with a sufficient number of L2 participants who could eventually constitute a statistically sound sample.

#### *3.2.2.1.3 Exposure to CA*

Recall that the IPG participants received input in the MSA and EA varieties due to the structural nature of their program, which aimed to simultaneously introduce both varieties from the outset and throughout. In contrast, the UIPG participants received Arabic input in MSA only due to the structural nature of their program, which aimed to introduce the MSA variety from the outset and

throughout. However, because CA materials are integrated in the textbook of the UIPG and because CA is commonly presented in TV and used by native speakers of Arabic ubiquitously, it was crucial to control for the variable related to the UIPG participants' exposure to CA. Thus, to recruit only L2ers representing the UIPG who were homogenous in terms of receiving input in MSA only, the following two questions were incorporated into the LHQ: (a) Do you self-study the sections designed for *Al- 'āmmiyya* "Arabic dialects" in your textbook? and (b) Do you get exposed to any Arabic dialects outside the classroom (e.g., communication with friends, watching TV)? Those who answered positively to either one of the two questions were asked to state the duration of their self-study of CA and the duration of their exposure to CA through their Arabic friends or TV. The choices provided in the LHQ were as follows: 1–3 hours per week, 4–6 hours per week, and more than 6 hours per week. Moreover, the LHQ contained the EA sentence in (1) for which the participants were asked to report whether they had a full comprehension of it. If so, they were further asked to provide an English translation for it, without using any dictionary.

- (1) *ʔil-bint*                      *ma-b-tā-kul-š*                      *kuwayyis*                      *zayy*                      *kull*  
the-girl                      Neg-Prog-Imperf.3.s.f-eat-Neg                      well                      like                      every  
*yōm*                      *ʕašān*                      *ʕayyān-a*                      *ʔin-nahār*                      *dah*  
day                      because                      sick-s.f                      the-day.m                      this.m  
"The girl is not eating well like every day because she is sick today."

Accordingly, three types of L2ers were excluded from the recruitment of the UIPG: those who reported self-study of CA for more than 3 hours per week, those who reported exposure to CA through TV or friends for more than 6 hours per week, and those who provided entirely accurate translations of the provided EA sentence.

#### 3.2.2.1.4 Code-switching behavior

Another essential criterion considered in the recruitment of L2 participants pertained to the code-switching behavior between MSA and CA. To proceed with the study with a certain degree of assurance that the IPG participants would produce data in both MSA and EA while their UIPG counterparts would produce data in MSA only, they were asked to indicate the frequency of code-switching behavior between MSA and CA in their spoken discourse on a Likert point scale: never, rarely, sometimes, regularly, often, usually, and always. Accordingly, those in the IPG who reported they *never* code-switched were excluded, and those in the UIPG who reported they *often, usually, or always* code-switched were also excluded. The results of the code-switching frequency among the IPG and UIPG are discussed in Section 3.2.2.2.5 below.

#### 3.2.2.1.5 Proficiency level

Another critical criterion taken into account in the recruitment process of L2 participants pertained to their proficiency levels. Precise determination of the proficiency level of each L2 participant was a necessary step for increasing the level of homogeneity and compatibility between all the participants in their respective proficiency levels—first year, second year, and third year—and hence to achieve valid results that were based on statistically sound data. Therefore, this step entailed adopting three measures that were expected to correctly assess the participants' proficiency-placement level and ensure accurate placement of all of them in their respective subgroups of proficiency.

The first indicator of proficiency-placement was the level of Arabic classes in which they were enrolled during the time of recruitment. However, exclusive reliance on this measure was insufficient due to the possibility of misplacing the participants in inaccurate and irrelevant proficiency levels in their programs. Therefore, a second layer of measures was implemented. It

entailed asking the L2 participants to self-rate their Arabic proficiency based on a Likert point scale: very weak, weak, limited, average, good, very good, and excellent.

To reach the maximum level of assurance that the IPG and UIPG at each proficiency level were comparably homogenous in terms of their proficiency-placement levels in their Arabic programs, a third layer of measures was implemented. It entailed taking a version of the Arabic placement test on reading comprehension, which is used in the Arabic Linguistics Institute at King Saud University. This test involved six texts varying in length. Each text was followed by several questions (13 questions in total), each of which was followed by four choices. The reason for choosing the test on reading comprehension, among other skills, was that unlike the speaking and listening skills that can be taught and presented in different varieties of Arabic, reading materials are typically presented in MSA only, regardless of the structural and instructional nature of Arabic L2 programs. Hence, all Arabic L2ers are expected to be exposed to reading input in the MSA variety only, irrespective of the medium of instruction adopted by their programs. The results of the second and third proficiency measures are discussed in Section 3.2.2.2.6 below.

### ***3.2.2.2 L2 Participants recruitment results***

#### *3.2.2.2.1 Demographic information of the L2 participants*

Based on the recruitment criteria discussed above in Section 3.2.2.1, a total of 57 L2 participants representing the IPG and UIPG were recruited. Table 3.2 displays the demographic information of the L2 participants. The IPG participants were learning MSA and EA simultaneously at all three proficiency levels, whereas their UIPG counterparts were learning MSA only from the outset and throughout. Overall, 43 of them reported some knowledge in L2s other than Arabic.

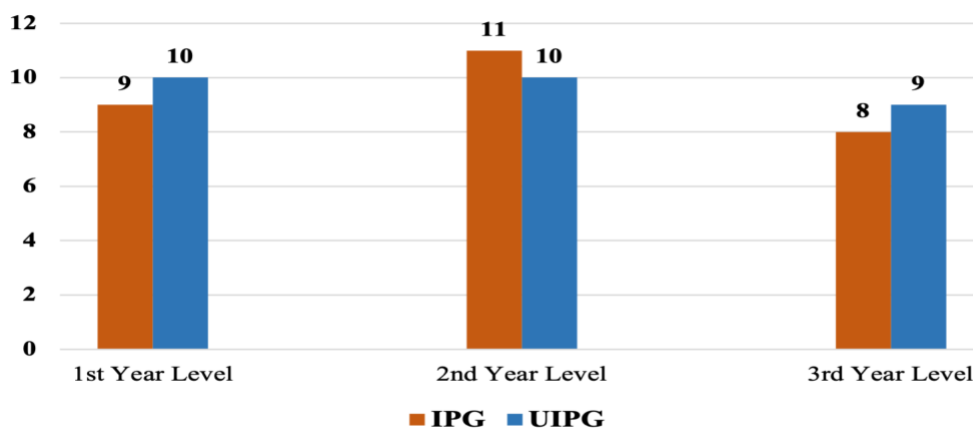
**Table 3.2** General Demographic Information of the IPG and UIPG at Each Proficiency Level

Group and Proficiency Level	Gender M/F/NB	Age Range	Age Mean	Other L2s Learned	Variety Learned
<b>IPG</b>					
1st year ( <i>n</i> = 9)	2/7/0	18–20	18.9	8	MSA & EA
2nd year ( <i>n</i> = 11)	4/7/0	19–20	19.1	7	MSA & EA
3rd year ( <i>n</i> = 8)	4/4/0	20–21	20.3	6	MSA & EA
<b>UIPG</b>					
1st year ( <i>n</i> = 10)	3/6/1	18–20	18.9	9	MSA
2nd year ( <i>n</i> = 10)	4/6/0	19–28	20.4	7	MSA
3rd year ( <i>n</i> = 9)	5/4/0	20–22	20.3	6	MSA

Note. M/F/NB = males/females/non-binary.

Regarding the number distribution, the 57 L2 participants belonged comparably to the two L2 main groups. Figure 3.1 shows the number distribution of the L2 participants in the IPG and UIPG at each proficiency level. Overall, 28 participants represented the IPG: 9 first-year participants, 11 second-year participants, and 8 third-year participants. As for the remaining 29 L2 participants, they represented the UIPG: 10 first-year participants, 10 second-year participants, and 9 third-year participants.

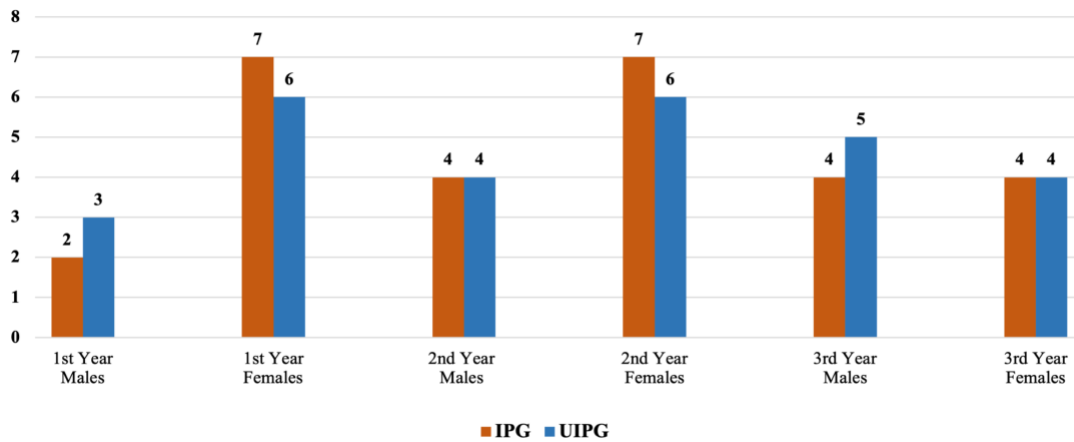
**Figure 3.1** Number Distribution of the IPG and UIPG at Each Proficiency Level



Regarding gender, 22 males, 34 females, and 1 non-binary individual took part in the study. Among the 28 participants in the IPG, only 10 were males while the remaining 18 were females. Their gender distribution across the three proficiency levels was as follows: 2 males and 7 females were first-year L2ers; 4 males and 7 females were second-year L2ers; 4 males and 4 females were third-year L2ers.

As for the 29 participants in UIPG, 12 were males, 16 were females, and 1 was nonbinary. Their distribution across the three proficiency levels was as follows: 3 males, 6 females, and 1 nonbinary were first-year L2ers; 4 males and 6 females were second-year L2ers; 5 males and 4 females were third-year L2ers. Figure 3.2 shows the gender distribution of the L2 participants in the IPG and UIPG across all proficiency levels.

**Figure 3.2** Gender Distribution of the IPG and UIPG at Each Proficiency Level



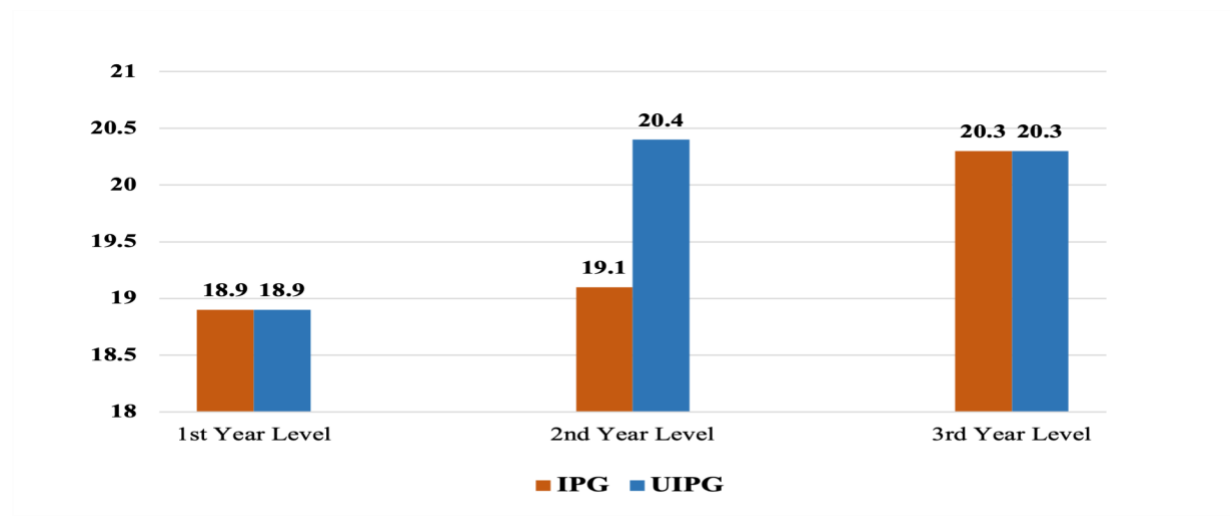
*Note.* The 10th first-year participant in the UIPG identified their gender as non-binary.

Regarding age, the L2 participants fell in the range of 18–28 years old. The age distribution of the IPG participants was as follows: the first-year participants were in the range of 18–20 years, with an age mean of 18.9 years; the second-year participants were in the range of

19–20 years, with an age mean of 19.1 years; the third-year participants were in the range of 20–21 years, with an age mean of 20.3 years.

As for the age distribution of the UIPG participants, it was as follows: the first-year participants were in the range of 18–20 years, with an age mean of 18.9 years; the second-year participants were in the range of 19–28 years, with an age mean of 20.4 years; the third-year participants were in the range of 20–22 years, with an age mean of 20.3 years. Figure 3.3 shows the age mean of the L2 participants in the IPG and UIPG across all proficiency levels.

**Figure 3.3** Age Means of the IPG and UIPG at Each Proficiency Level



As for the education level, the majority of the L2 participants were college students during the time of data collection. Precisely, all 28 participants representing the three proficiency levels of the IPG were college students, as were 27 participants in the UIPG. For the remaining two UIPG participants, one was an associate degree student, whereas the other was a master’s student, and both were second-year participants.



#### 3.2.2.2.2 *L1 background of the L2 participants*

All 57 participants reported English as their L1, and none of them was a heritage speaker of Arabic. However, it is worth mentioning that a number of L2ers in the UIP were of Islamic heritage. The inclusion and exclusion of those L2ers in the study were based on their performance compared to their peers in their respective subgroup (i.e., proficiency level). To illustrate, Islamic heritage L2ers whose performance was within or below the mean of their peers in their respective subgroups were included in the study. However, those whose performance was significantly above the mean were excluded. The exclusion of the latter was due to their likely intensive exposure to Arabic through religious sources (e.g., *Qur'ān*), which have most likely provided them with an input advantage over their peers of non-Islamic heritage—and thus they were excluded from the study. Ultimately, only two L2ers of Islamic heritage were included in the study: one was a first-year participant and the other was a second-year participant, and both were in the UIPG.

#### 3.2.2.2.3 *Prior exposure to Arabic*

All 57 L2 participants were comparably homogenous in their prior exposure to Arabic. That is, none of them had resided in any Arabic-speaking countries for more than 3 months, and none had studied Arabic in previous programs for more than 6 months. Precisely, only two participants had studied Arabic in previous programs for no more than 3 months: one first-year participant in the UIPG and one second-year participant in the IPG. Seven participants had previously studied Arabic for a duration up to 6 months: two first-year participants in the IPG, two first-year participants in the UIPG, one second-year participant in the UIPG, and two third-year participants in the IPG.

### 3.2.2.2.4 Exposure to CA

The responses collected from the L2 participants on the question related to their self-study of CA sections in their textbooks revealed that, contrary to the IPG participants, most of whom reported self-study of CA materials in their textbook, only one second-year UIPG participant self-studied the CA sections for only 1–3 hours per week. Regarding the question on the time duration of exposure to CA input through TV or communication with friends, 12 UIPG participants had such exposure for 1–3 hours per week and four participants for 4–6 hours per week. Comparably, 14 IPG participants reported exposure to CA for 1–3 hours a week and five participants for 4–6 hours a week. Table 3.3 displays the time duration of exposure to CA among the participants in the IPG and UIPG at each proficiency level.

**Table 3.3** Time Duration of Exposure to CA Input among the IPG and UIPG Across All Proficiency Levels

Group and Proficiency Level	Self-Study of CA (1–3 h/w)	Self-Study of CA (4–6 h/w)	Self-Study of CA (+ 6 h/w)	Ex to CA via TV/ Friends (1–3 h/w)	Ex to CA via TV/Friends (4–6 h/w)
<b>IPG</b>					
1st year ( $n = 9$ )	7/9	1/9	1/9	4/9	2/9
2nd year ( $n = 11$ )	4/11	4/11	1/11	8/11	2/11
3rd year ( $n = 8$ )	3/8	3/8	0/8	2/8	1/8
<b>UIPG</b>					
1st year ( $n = 10$ )	0/10	0/10	0/10	2/10	2/10
2nd year ( $n = 10$ )	1/10	0/10	0/10	7/10	1/10
3rd year ( $n = 9$ )	0/9	0/9	0/9	3/9	1/9

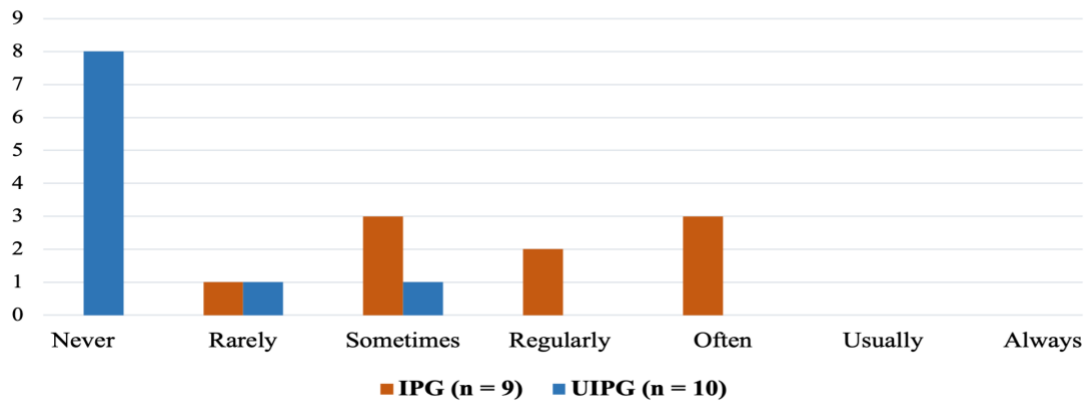
*Note.* Ex = exposure; h/w = hours per week.

Regarding the translation question, only one first-year participant of the 29 UIPG participants reported some understanding of the EA sentence provided in the LHQ. However, the translation was inaccurate. Conversely, 20 IPG participants provided translations to the EA sentence: six first-year participants, eight second-year participants, and six third-year participants. Only three were entirely accurate, whereas the other 17 were partially accurate.

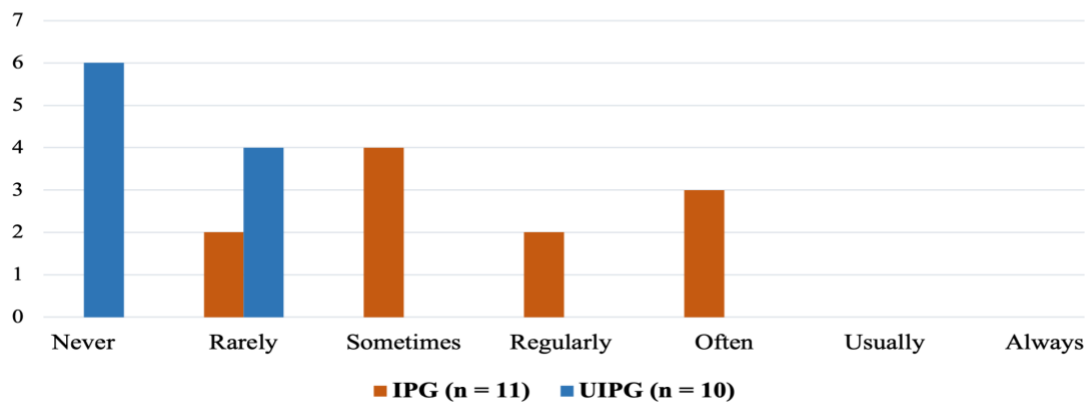
### 3.2.2.2.5 Code-switching behavior

The responses on the question related to the code-switching frequency varied across all participants of both L2 groups, as displayed in Figures 3.4, 3.5, and 3.6 for the first-year, second-year, and third-year participants, respectively. In general, most of the IPG participants reported that they *sometimes*, *regularly*, and *often* code-switched between MSA and CA. Conversely, the majority of the UIPG participants reported that they *never* and *rarely* code-switched.

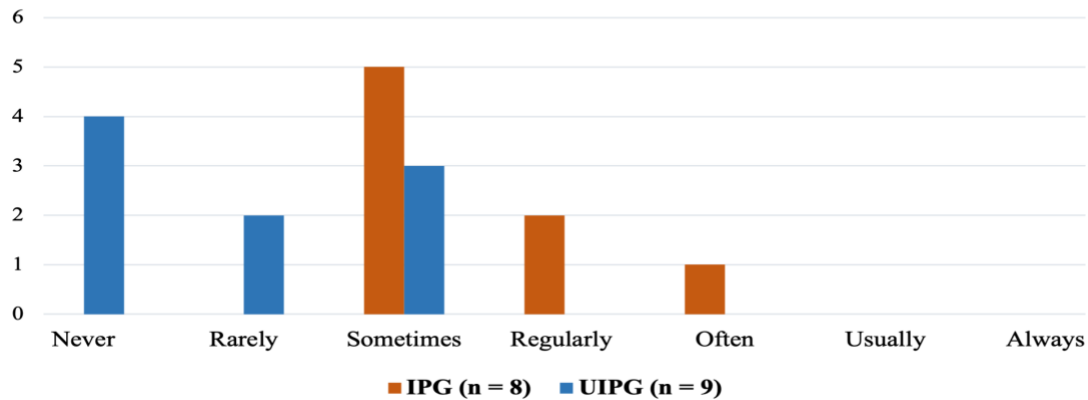
**Figure 3.4** Code-Switching Frequency among First-Year Participants in the IPG and UIPG



**Figure 3.5** Code-Switching Frequency among Second-Year Participants in the IPG and UIPG



**Figure 3.6** Code-Switching Frequency among Third-Year Participants in the IPG and UIPG



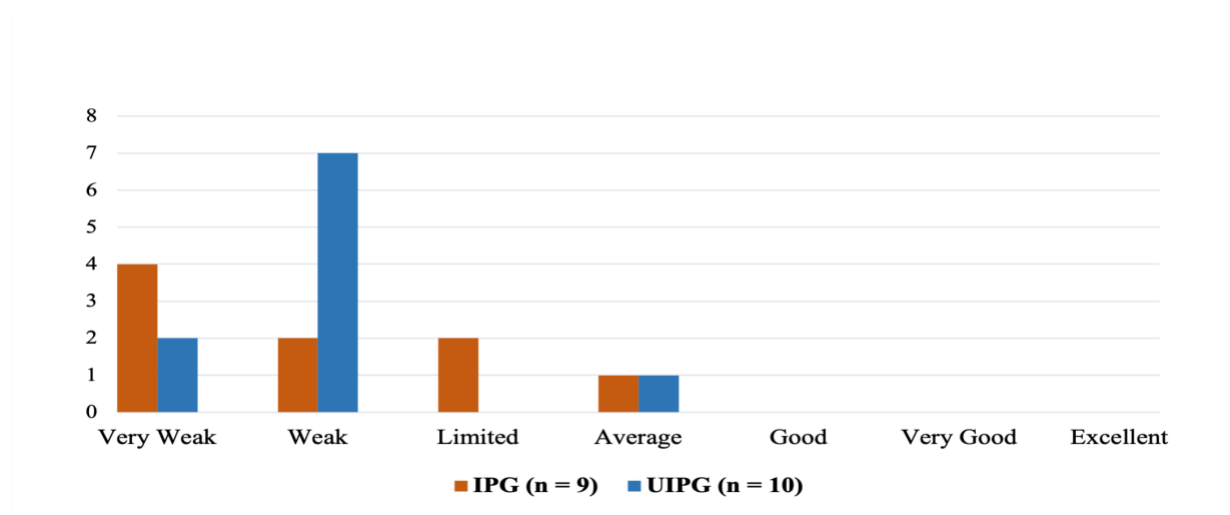
To examine the hypothesis that the code-switching frequency was significantly higher in the spoken discourse of the IPG compared to that of the UIPG at all proficiency levels, three Mann Whitney *U*-tests were performed separately in SPSS (one for each proficiency level). The frequency responses were entered as ordinal value numbers, where the value of “never” was 1 and the value of “always” was 7. An alpha of .05 was utilized as an indicator for the significance level. The three Mann Whitney *U*-tests revealed the following: the first-year IPG participants were associated with a significantly larger median of code-switching frequency ( $Md = 3.00, n = 9$ ) compared to the first-year UIPG participants ( $Md = 1.00, n = 10$ ),  $U = 6.50, z = -3.28, p = .001$ , with a great effect size ( $r = .75$ ); the second-year IPG participants were likewise associated with a significantly larger median of code-switching frequency ( $Md = 3.00, n = 11$ ) compared to the second-year UIPG participants ( $Md = 1.00, n = 10$ ),  $U = 4.00, z = -3.69, p = .001$ , with a great effect size ( $r = .81$ ); the third-year IPG participants were also associated with a significantly larger median of code-switching frequency ( $Md = 3.00, n = 8$ ) compared to the third-year UIPG participants ( $Md = 2.00, n = 9$ ),  $U = 7.50, z = -2.92, p = .004$ , with a great effect size ( $r = .71$ ). Thus, the hypothesis that the code-switching frequency in the spoken discourse of

the IPG was at a significantly higher rate compared to the UIPG at each proficiency level was confirmed. Accordingly, the current study was carried out with a certain degree of confidence that the IPG participants would produce data in both MSA and EA, whereas their UIPG counterparts would produce data in MSA only.

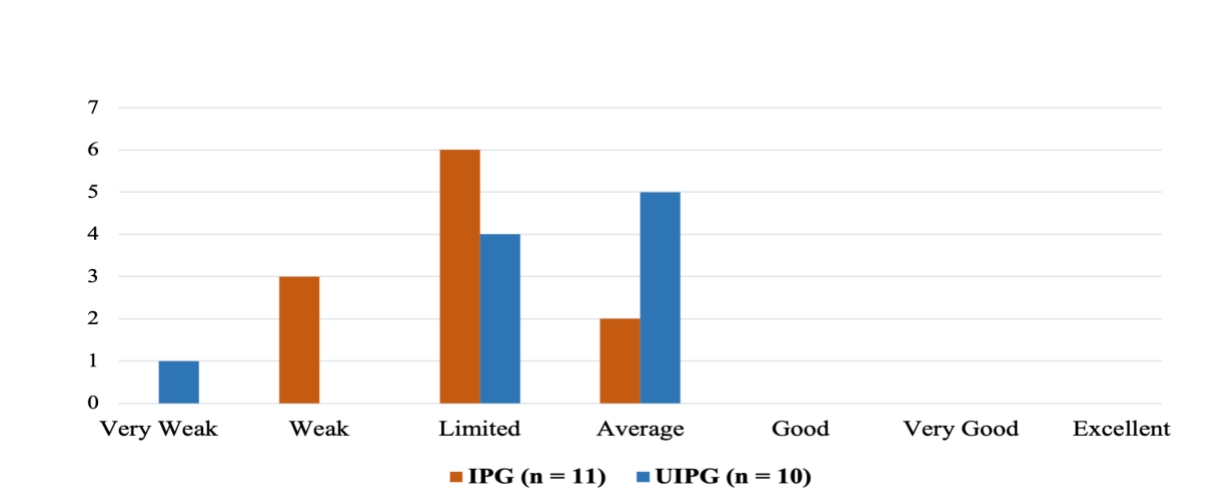
### 3.2.2.2.6 Proficiency level

As discussed in Section 3.2.2.1.5 above, three different layers of measures were used to precisely determine whether the participants in the IPG and UIPG were proficiently homogenous at each of the three proficiency levels. That is, besides the current levels of Arabic classes in which the L2 participants were enrolled, they were asked to rate their Arabic proficiency and take an Arabic reading proficiency test. The obtained responses on the self-rating of Arabic proficiency varied across all participants in both L2 groups, as displayed in Figures 3.7, 3.8, and 3.9 for the first-year, second-year, and third-year participants, respectively.

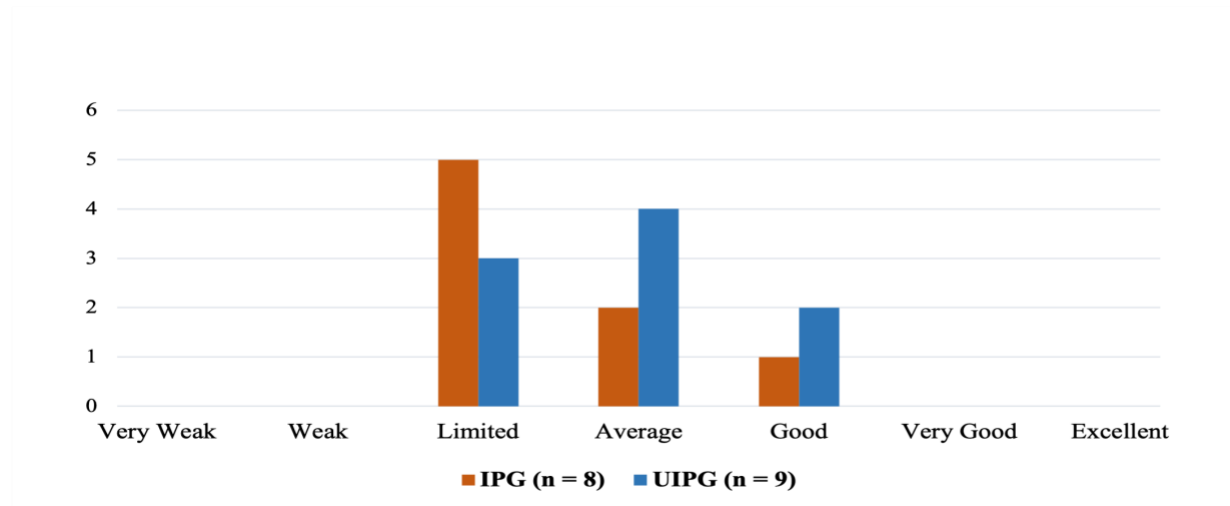
**Figure 3.7 Proficiency Self-Rating Responses by First-Year Participants in the IPG and UIPG**



**Figure 3.8** Proficiency Self-Rating Responses by Second-Year Participants in the IPG and UIPG



**Figure 3.9** Proficiency Self-Rating Responses by Third-Year Participants in the IPG and UIPG



To examine the hypothesis that the IPG and UIPG at each of the three proficiency level—first year, second year, and third year—were homogenous and compatible in their self-perceptions of their Arabic proficiency, three Mann Whitney *U*-tests were performed separately in SPSS (one for each proficiency level). The frequency responses were entered as ordinal value numbers, where the value of “very weak” was 1 and the value of “excellent” was 7. An alpha of .05 was utilized as an indicator for the significance level. The three Mann Whitney *U*-tests

revealed the following: there was no significant difference between the first-year IPG participants ( $Md = 2.00, n = 9$ ) and their UIPG counterparts ( $Md = 2.00, n = 10$ ),  $U = 42.50, z = -.22, p = .826$ , with a medium effect size ( $r = .05$ ); there was no significant difference between the second-year IPG participants ( $Md = 3.00, n = 11$ ) and their UIPG counterparts ( $Md = 3.50, n = 10$ ),  $U = 36.00, z = -1.45, p = .148$ , with a medium effect size ( $r = .32$ ); there was no significant difference between the third-year IPG participants ( $Md = 3.00, n = 8$ ) and their UIPG counterparts ( $Md = 4.00, n = 9$ ),  $U = 25.50, z = -2.10, p = .273$ , with a small effect size ( $r = .27$ ). Thus, both the IPG and UIPG were comparably homogenous in their self-perceptions of their Arabic proficiency at each proficiency level.

As for the Arabic reading proficiency-placement test, the data obtained from the L2 participants on the test were converted into percentages of correct answers. To examine the hypothesis that the IPG and UIPG at each of the three proficiency levels were statistically homogenous in their performance on the proficiency-placement test, three independent samples  $t$ -tests were performed separately in SPSS (one for each proficiency level). The data obtained from the proficiency test were entered as continuous values. An alpha of .05 was utilized as an indicator for the significance level. For the first-year level, the assumption for equality of variances was assessed and satisfied via Levene's  $F$  test, ( $F(17) = .502, p = .488$ ). The analysis revealed that the first-year IPG participants ( $n = 9$ ) were associated with a smaller score mean ( $M = 21.38, SD = 18.75$ ), compared to their UIPG counterparts ( $n = 10$ ), who were associated with a greater score mean ( $M = 24.64, SD = 15.31$ ). However, this difference between the two score means was not significant ( $t(17) = -.417, p = .682$ ). For the second-year level, the assumption for equality of variances was assessed and satisfied via Levene's  $F$  test ( $F(19) = .945, p = .343$ ). The analysis revealed that the second-year IPG participants ( $n = 11$ ) were associated with a smaller

score mean ( $n = 23.10$ ,  $SD = 9.74$ ), compared to their UIPG counterparts ( $n = 10$ ), who were associated with a greater score mean ( $M = 27.72$ ,  $SD = 10.39$ ). However, this difference between the two score means was not significant ( $t(19) = -1.05$ ,  $p = .306$ ). For the third-year level, the assumption for equality of variances was assessed and satisfied via Levene's  $F$  test ( $F(15) = .952$ ,  $p = .570$ ). The analysis revealed that the third-year IPG participants ( $n = 8$ ) were associated with a smaller score mean ( $M = 31.75$ ,  $SD = 10.41$ ), compared to their UIPG counterparts ( $n = 9$ ), who were associated with a greater score mean ( $M = 43.61$ ,  $SD = 14.37$ ). However, this difference between the two score means was not significant ( $t(15) = -1.925$ ,  $p = .073$ ). Thus, both the IPG and UIPG at each proficiency level were comparably homogenous in their performance on the proficiency-placement test.

Hence, the two measures of proficiency self-rating and proficiency-placement test confirmed the hypothesis of great homogeneity and compatibility between the IPG and UIPG at each proficiency level. This great homogeneity provided, in turn, a high degree of certainty that all the participants were correctly placed at their respective proficiency levels before the data analysis was carried out.

### **3.2.3 Control participants**

Besides the two L2 groups (IPG and UIPG), a CG representing native speakers of EA was recruited. The inclusion of the controls served two essential purposes. First, it helped overcome one of the major limitations of many studies investigating the fluency, accuracy, and complexity triad of L2ers (Skehan, 2009). Second, to control for the variable related to the code-switching/mixing behavior between MSA and EA in the production of the L2 participants, the production data obtained from the control participants functioned as a reference to decide whether those instances produced by the L2 participants were authentic (i.e., code-switching) or



unauthentic (i.e., code-mixing). Thus, the performance of the controls established a baseline along which the proficiency triad of the L2 participants and the production instances of code-switching/mixing between MSA and EA were compared and measured.

The control participants were asked to fill out a modified and short version of the LHQ distributed to the L2 participants. The control LHQ (see Appendix B) involved questions related to their gender, age, L1, education level, field of study, time length in the United States, and code-switching frequency in their spoken discourse. The most critical variables that were considered when recruiting the controls were their L1, length of stay in the United States, field of study, and education level. The recruitment criteria and results of the control participants are discussed below in Sections 3.2.3.1 and 3.2.3.2, respectively.

### ***3.2.3.1 Control participant recruitment criteria***

The L1 of the controls had to be EA. Additionally, because length of stay in foreign countries may cause attrition in one's L1, only those who stayed in the United States for less than 5 years were included in the recruitment. Furthermore, to eliminate any possibilities of metalinguistic familiarity and conscious grammatical knowledge of the target structures, the study field of the controls had to be in different areas other than Arabic linguistics and Arabic literature.

Accordingly, the following three types of participants were excluded from the recruitment: those who reported languages or Arabic varieties other than EA as their L1, those who lived in the United States for more than 5 years, and those who majored in Arabic linguistics or Arabic literature. Moreover, since the study aimed to examine how the target structures in both varieties of MSA and EA were produced by the IPG, it was critical to recruit educated native speakers. This is so because many studies have revealed that features that are exclusively used in MSA were not prevalent in the discourse of uneducated native Arabic

speakers, as opposed to educated native Arabic speakers whose discourse extensively exhibited features from MSA and CA at the same time (e.g., Al Aslaa & Alhawary, 2021; Albirini, 2011; Badawi, 1973; Chakrani, 2015; Ryding, 1991). Accordingly, only native EA speakers who were pursuing graduate education were included in the CG recruitment.

### 3.2.3.2 Control participant recruitment results

Based on the control participant recruitment criteria discussed above, a total of nine native speakers of EA were recruited. Table 3.4 displays the demographic information of the control participants. Three were females, and six were males. They were in the range of 28–38 years old, with an age mean of 32.3 years. The length of their stay in the United States was in the range of 1–5 years, with a mean of 3.4 years. They all specialized in majors unrelated to Arabic and were pursuing doctoral and postdoctoral degrees at a prestigious university in the United States, and they all provided accurate translations of the EA sentence included in the control LHQ.

**Table 3.4** Demographic Information of the Control Participants

Controls EA L1	Gender	Age	Length of Stay in the U.S.	Major	Level of Education
1	F	35	1 year	Pediatric Dentistry	Doctorate
2	F	30	3 years	Pharmacy	Doctorate
3	M	29	3 years	Pharmacy	Doctorate
4	M	29	3 years	Electrical Engineering	Doctorate
5	M	28	1 year	Political Science	Doctorate
6	F	31	2 years	Geology	Doctorate
7	M	38	5 years	Biochemistry	Post-doctorate
8	M	33	2 years	Neurology	Post-doctorate
9	M	38	1 year	Cancer Biology	Post-doctorate

*Note.* F = female; M = male.

### 3.3 Target Structures

This section provides a detailed description of the target structures of the simple past tense/aspect, simple present tense/aspect, past negation, verbal present negation, subjunctive

subordination, and cause–effect subordination in the varieties of MSA and EA. All provided descriptions and examples are intended to focus exclusively on the data-related features examined in the current study. Marking/ending suffixes for the imperfect verb mood (indicative, subjunctive, and jussive) and for the noun and adjective case (nominative, accusative, and genitive) were not within the scope of the study. Therefore, they are provided in parentheses in all MSA examples provided here to indicate their optionality use by all participants. It is worth mentioning that the past/perfect verb does not exhibit any ending mood and that the mood and case ending suffixes are not overtly spelled out in EA and other CA varieties. Thus, ending suffixes for mood and case are not included in all EA examples.

The description of the target structures in MSA was mainly based on contemporary Arabic grammar books (e.g., Alhawary, 2011) and L2 participants' textbooks, which are Parts I and II of the second and third edition of *Al-Kitāb fī Ta'allum Al-'Arabiyya* (Brustad et al., 2004, 2007, 2011, 2013). As for the description of the target structures in EA, it was based on the third edition of textbook, as well as the CG's production data.

### **3.3.1 Tense/Aspect**

Focus here is on two tense/aspect structures: simple past tense/aspect and simple present tense/aspect. Before discussing these two structures, a brief review of three relevant issues is in order: types of Arabic sentences, the Arabic verbal agreement system, and perspectives of grammarians on the temporal and aspectual information conveyed by Arabic verbs.

Sentences in all Arabic varieties may be classified into two types: verbal and nominal. The former refers to sentences in the VSO word order. As for the latter, it is divided into two sentential subtypes: nominal-verbless sentences whose heads are nouns or adjectives, and nominal-verbal sentences whose heads are verb predicates. The latter subtype refers to sentences

in the SVO word order (e.g., Albirini et al., 2019; Alhawary, 2009a, 2011, 2019; Aoun et al., 2010). In the current study, focus is on verbal sentences (i.e., VSO) and nominal-verbal sentences (i.e., SVO) only.

Although verbal agreement is not within the main focus of the current study, it is strongly relevant because the production data were elicited in the contexts of third-person singular masculine (referred henceforth as 3.s.m) and third-person singular feminine (referred henceforth as 3.s.f). The phi/agreement inflectional features for person, gender, and number are checked/realized asymmetrically in Arabic, depending on the sentence word order: SVO and VSO. While the former entails full agreement between the verb and its preverbal subject in all agreement features, the latter entails partial agreement between the verb and its postverbal subject in the gender feature only. However, this asymmetry in the inflectional agreement features between the two types of word order has no impact on this study because all production collected data were carried out in the 3.s.m and 3.s.f contexts, which are inflected on verb forms symmetrically in both word orders (see also Alhawary 2009a, 2011, 2019). To elaborate, on simple past/perfect verbs, the agreement features of person, gender, and number are inflected as a zero suffix (i.e., null morpheme) for the 3.s.m form, as in (2a), and as the suffix *-at* for the 3.s.f form, as in (2b). Conversely, on simple present/imperfect verbs, these three agreement features are inflectionally conflated in the prefix *ya/ya-* for the 3.s.m form and in the prefix *ta/tu-* for the 3.s.f form, as in (3a) and (3b), respectively. This inflection of agreement features behaves identically in both types of word orders. Accordingly, the examples in (2a–b) and (3a–b) that illustrate, respectively, the inflectional agreement features on simple past tense/aspect verb and simple present tense/aspect verbs for the 3.s.m and 3.s.f forms in the SVO word order are identical to those in (4a–b) and (5a–b) that illustrate, respectively, the inflectional agreement

features on simple past tense/aspect verbs and simple present tense/aspect verbs for the 3.s.m and 3.s.f forms in the VSO word order.

- (2) a. *ʔal-walad(-u)*                      *ʔakala*  
the-boy-(Nom)                      Perf.ate.3.s.m  
“The boy ate.”
- b. *ʔal-bint(-u)*                      *ʔakal-at*  
the-girl-(Nom)                      Perf.ate-3.s.f  
“The girl ate.”
- (3) a. *ʔal-walad(-u)*                      *ya-ʔkul(-u)*  
the-boy-(Nom)                      Imperf.3.s.m-eat(-Indic)  
“The boy eats/is eating.”
- b. *ʔal-bint(-u)*                      *ta-ʔkul(-u)*  
the-girl-(Nom)                      Imperf.3.s.f-eat(-Indic)  
“The girl eats/is eating.”
- (4) a. *ʔakala*                      *ʔal-walad(-u)*  
Perf.ate.3.s.m                      the-boy(-Nom)  
“The boy ate.”
- b. *ʔakal-at*                      *ʔal-bint(-u)*  
Perf.ate-3.s.f                      the-girl-(Nom)  
“The girl ate.”
- (5) a. *ya-ʔkul(-u)*                      *ʔal-walad(-u)*  
Imperf.3.s.m-eat(-Indic)                      the-boy(-Nom)  
“The boy eats/is eating.”
- b. *ta-ʔkul(-u)*                      *ʔal-bint(-u)*  
Imperf.3.s.f-eat(-Indic)                      the-girl-(Nom)  
“The girl eats/is eating.”

Another important issue pertains to the temporal and aspectual information conveyed by the verb forms. Like the view held by early Arab grammarians, most contemporary Arabic linguists argue

that Arabic verbs mark both *tense* and *aspect* (e.g., Abboud & McCarus, 1983; Alhawary, 2009a, 2011, 2019; Fassi Fehri, 1993; McCarthy, 1979). Thus, these linguists usually use the term *perfect past* to refer to any action that took place in the past and the term *imperfect present* to refer to any incomplete action. They rely on two arguments to support their stance. First, the vocalic melody of the verb indicates the temporal and aspectual information. That is, the form *faʕala*, which has the CvCvCv pattern, expresses a finished event or action that took place in the past, and the form *yafʕalu*, which has the CvCCvCv pattern, expresses a present/unfinished action (McCarthy, 1979).<sup>7</sup>

Second, Arabic is a morphologically rich language that implements various mechanisms, such as the phi/agreement features that are attached to the verb to minimize the occurrences of structural ambiguity (Shormani, 2014) and to evoke the temporal and aspectual information (Alhawary, 2009a, 2011, 2019). Hence, besides the vocalic melody of verbs, the agreement features convey the temporal and aspectual information. Fassi Fehri (1993) advanced this argument by remarking that the verb with these agreement features must be compatible with the temporal adverbs, as in (6a). Otherwise, the whole structure crashes, as in (6b).

- |     |    |  |                                   |                          |
|-----|----|--|-----------------------------------|--------------------------|
| (6) | a. | <i>ʔal-bint(-u)</i><br>the-girl(-Nom)<br>“The girl ate yesterday.” | <i>ʔakal-at</i><br>Perf.ate-3.s.f | <i>ʔams</i><br>yesterday |
|     | b. | <i>*ʔal-bint(-u)</i><br>the-girl(-Nom)<br>“The girl ate tomorrow.” | <i>ʔakal-at</i><br>Perf.ate-3.s.f | <i>yadan</i><br>tomorrow |

---

<sup>7</sup> These two verb patterns are not the only common verb forms in Arabic. There are nine other common forms for the past and present tenses, and all of them exhibit predictable patterns (e.g., Alhawary, 2011).

Therefore, the temporal information (past, present, and future) and aspectual information (perfect and imperfect) in MSA is expressed “paraphrastically, that is, by means of one of two forms of the verb and additional simple words” (Alhawary 2011, p. 80). This also applies to all Arabic varieties, including EA—the variety under study.<sup>8</sup>

### 3.3.1.1 Past tense/aspect

MSA and EA exhibit only one morphosyntactic similarity when it comes to conveying aspectual information. This similarity manifests itself in expressing the simple past/perfect tense/aspect (referred henceforth as the past tense). To illustrate, similar to English, the past tense is expressed in both Arabic varieties (MSA and EA) by the verb perfect form and usually with an adverb indicating a past time, as in (7a–b) and (8a–b) for MSA and EA, respectively.

Notwithstanding the phonological and lexical variations between MSA and EA, both varieties do not require any aspectual markers before the perfect form of the verb, which must be in full agreement with the subject.

- (7) a. *ʔal-walad(-u)*      *ʕamila*                      *ʔal-wājib(-a)*                      *ʔams*  
the-boy(-Nom)      Perf.did.3.s.m                      the-homework(-Acc)                      yesterday  
“The boy did the homework yesterday.”
- b. *ʔal-bint(-u)*              *ʕamil-at*                      *ʔal-wājib(-a)*                      *ʔams*  
the-girl(-Nom)      Perf.did-3.s.f                      the-homework-(Acc)                      yesterday  
“The girl did the homework yesterday.”

---

<sup>8</sup> Horesh (2008) examined the issue of tense among native speakers of different Arabic dialects. The participants were provided with some sentences and were asked to determine whether those utterances had past, present, or future interpretations. The participants’ responses were consistent because all of them constantly interpreted verbs with suffixes as completed actions that took place in the past, whereas verbs with prefixes and suffixes did not convey any information indicating that the action had taken place.

- (8) a. *ʔil-walad*      *ʕimil*      *ʔil-wāgib*      *ʔams*  
the-boy      Perf.did.3.s.m      the-homework      yesterday  
“The boy did the homework yesterday.”
- b. *ʔil-bint*      *ʕiml-at*      *ʔil-wāgib*      *ʔams*  
the-girl      Perf.did-3.s.f      the-homework      yesterday  
“The girl did the homework yesterday.”

Although the past progressive tense/aspect (referred henceforth as the past progressive) was not a target structure in the current study, it was present in the data collected from some participants, including the controls. To this effect, any production of this structure was tolerated when coding the data and counted toward the total production of the past tense. The past progressive behaves similarly in MSA and English. That is, it is expressed in MSA by means of placing the modal *kāna* “was/were,” which has to be in full agreement with the subject, before the imperfect verb and by using an adverb of time indicating the past tense, as in (9a–b). However, this tense is conveyed in EA by means of placing the progressive modal *kana* “was/were” before the imperfect verb, which has to be attached to the progressive prefix *bi-*, as in (10a–b).

- (9) a. *ʔal-walad(-u)*      *kāna*      *ya-drus(-u)*      *ʔams*  
the-boy-(Nom)      Perf.was.3.s.m      Imperf.3.s.m-study(-Indic)      yesterday  
“The boy was studying yesterday.”
- b. *ʔal-bint(-u)*      *kān-at*      *ta-drus(-u)*      *ʔams*  
the-girl(-Nom)      Perf.was-3.s.f      Imperf.3.s.f-study(-Indic)      yesterday  
“The girl was studying yesterday.”
- (10) a. *ʔil-walad*      *kān*      *bi-y-zākir*      *ʔams*  
the-boy      Perf.was.3.s.m      Prog-Imperf.3.s.m-study      yesterday  
“The boy was studying yesterday.”
- b. *ʔil-bint*      *kān-it*      *bi-t-zākir*      *ʔams*  
the-girl      Perf.was-3.s.f      Prog-Imperf.3.s.f-study      yesterday  
“The girl was studying yesterday.”



### 3.3.1.2 Present tense/aspect

In the present temporal and aspectual distribution, MSA diverges from EA but converges with English. Similar to English, the simple present/imperfect tense/aspect (referred henceforth as the present tense) is expressed in MSA by the imperfect verb and usually with an adverb of time indicating the present tense, as in (11a–b). In contrast, EA utilizes the aspectual habituality marker *bi-* that is obligatorily attached as a prefix to the imperfect verb, as in (12a–b).

- |      |    |                                       |  |                                |                    |
|------|----|---------------------------------------|--|--------------------------------|--------------------|
| (11) | a. | <i>ʔal-walad(-u)</i><br>the-boy(-Nom) | <i>ya-drus(-u)</i><br>Imperf.3.s.m-study(-Indic) | <i>kull(-a)</i><br>every(-Acc) | <i>yawm</i><br>day |
|      |    | “The boy studies every day.”          |  |                                |                    |
|      | b. | <i>ʔal-bint(-u)</i><br>the-girl(-Nom) | <i>ta-drus(-u)</i><br>Imperf.3.s.f-study(-Indic) | <i>kull(-a)</i><br>every-(Acc) | <i>yawm</i><br>day |
|      |    | “The girl studies every day.”         |  |                                |                    |
| (12) | a. | <i>ʔil-walad</i><br>the-boy           | <i>bi-yi-zākir</i><br>Hb-Imperf.3.s.m-study      | <i>kull</i><br>every           | <i>yōm</i><br>day  |
|      |    | “The boy studies every day.”          |  |                                |                    |
|      | b. | <i>ʔil-bint</i><br>the-girl           | <i>bi-t-zākir</i><br>Hb-Imperf.3.s.f-study       | <i>kull</i><br>every           | <i>yōm</i><br>day  |
|      |    | “The girl studies every day.”         |  |                                |                    |

However, the production data of the control participants revealed that the aspectual marker of habituality *bi-* can be used optionally with a conjunctive verb in a conjoined clause (that is, after the conjunctions *wi* “and” and *ʔaw* “or”), as in (13a–b). It is also used optionally with a subordinate verb in a complex structure, as in (14a–b). The apparent reason for this optionality is that the temporal and aspectual information is already determined by/checked in the conjoined and matrix clauses, so the overt spell-out of the habitual marker *bi-*, without any semantic or pragmatic motive, would be redundant.

- (13) a. *ʔir-rāgil bi-yā-kul fiṭār-u wi yi-šrab ʔahwi-t-u*  
the-man Hb-Imperf.3.s.m-eat breakfast-his and Imperf.3.s.m-drink coffee-f-his  
“The man eats his breakfast and drinks his coffee.”
- b. *ʔis-sitt bi-tā-kul fiṭar-hā wi ti-šrab ʔahwi-t-hā*  
the-woman Hb-Imperf.3.s.f-eat breakfast-her and Imperf.3.s.f-drink coffee-f-her  
“The lady eats her breakfast and drinks her coffee.”
- (14) a. *ʔir-rāgil bi-yi-rūḥ li ʔin-nādī ʕašān yi-tmarran*  
the-man Hb-Imperf.3.s.m-go to the-gym because Imperf.3.s.m-exercise  
“The man goes to the gym [to] exercise.”
- b. *ʔis-sitt bi-t-rūḥ li ʔin-nādī ʕašān ti-tmarran*  
the-woman Hb-Imperf.3.s.f-go to the-gym because Imperf.3.s.f-exercise  
“The lady goes to the gym [to] exercise.”

### 3.3.2 Negation

#### 3.3.2.1 Past negation

MSA past sentences are negated by either the particle *lam* “did not” or the particle *mā* “did not.”

The latter is more marked and less prevalent because it is traditionally used for emphasis

(Alhawary, 2011). When the particle *lam* is used, the main verb must be in its imperfect form

and in the jussive mood, as in (15a–b), because this particle carries both the past tense and

negation. However, when the particle *mā* is used, the verb must be in its perfect form, as in (16a–

b). This is so because this particle does not carry any temporal information *per se* and is

considered a neutral negation marker that is traditionally used to negate verbal and nominal

sentences containing verb predicates in both temporal contexts of the past and present tense

(Alhawary, 2011; Fassi Fehri, 1993).

- (15) a. *ʔal-walad(-u) lam ya-drus*  
the-boy(-Nom) did not Imperf.3.s.m-study(Juss)  
“The boy did not study.”

- |      |    |  |                       |  |
|------|----|--|-----------------------|--|
|      | b. | <i>ʔal-bint(-u)</i><br>the-girl(-Nom)<br>“The girl did not study.” | <i>lam</i><br>did not | <i>ta-drus</i><br>Imperf.3.s.f-study(Juss) |
| (16) | a. | <i>ʔal-walad(-u)</i><br>the-boy(-Nom)<br>“The boy did not study.”  | <i>mā</i><br>Neg      | <i>darasa</i><br>Perf.study.3.s.m          |
|      | b. | <i>ʔal-bint(-u)</i><br>the-girl(-Nom)<br>“The girl did not study.” | <i>mā</i><br>Neg      | <i>daras-at</i><br>Perf.study-3.s.f        |

As for the EA past sentences, they are negated by the discontinuous particle *ma-š* that is realized as a proclitic and enclitic on perfect verbs, as in (17a – b).

- |      |    |                             |   |                           |
|------|----|-----------------------------|---|---------------------------|
| (17) | a. | <i>ʔil-walad</i><br>the-boy | <i>ma-zakir-š</i><br>Neg-Perf.study.3.s.m-Neg   | “The boy did not study.”  |
|      | b. | <i>ʔil-bint</i><br>the-girl | <i>ma-zakr-it-š</i><br>Neg-Perf.study-3.s.f-Neg | “The girl did not study.” |

### 3.3.2.2 Verbal present negation

As in English, MSA verbal present sentences, as well as nominal sentences containing verb predicates, are negated by the default negation particle *lā* “does/do not,” which heads a negation phrase above the main verbal phrase and does not cause any changes to the indicative mood of the imperfect form of verbs, as in (18a–b). Another negator is the particle *mā* “does/do not,” which behaves exactly as the particle *lā* “does/do not,” as in (19a–b). However, the optional particle *mā* is less prevalent because it is traditionally used to convey a semantically and programmatically emphasized negation construction (Alhawary, 2011). It should be noted that *mā* was not presented as a present negator in the L2ers’ textbooks. Thus, it was not counted to

toward the total correct production tokens of the verbal present negation structure, as will be discussed further in Section 3.6.2.4 below.

- (18) a. *ʔal-walad(-u)*                      *lā*                      *ya-drus(-u)*  
           the-boy(-Nom)                      does not                      Imperf.3.s.m-study(-Indic)  
           “The boy does not study.”
- b. *ʔal-bint(-u)*                      *lā*                      *ta-drus(-u)*  
           the-girl(-Nom)                      does not                      Imperf.3.s.f-study(-Indic)  
           “The girl does not study.”
- (19) a. *ʔal-walad(-u)*                      *mā*                      *ya-drus(-u)*  
           the-boy(-Nom)                      Neg                      Imperf.3.s.m-study(-Indic)  
           “The boy does not study.”
- b. *ʔal-bint(-u)*                      *mā*                      *ta-drus(-u)*  
           the-girl(-Nom)                      Neg                      Imperf.3.s.f-study(-Indic)  
           “The girl does not study.”

As for the EA verbal present sentences and nominal sentences containing verb predicates, they are typically negated by the EA discontinuous particle, which consists of two morphemes that are realized as the proclitic *ma-* and the enclitic *-š* on the verb predicate being negated, as in (20a–b). Another optional particle used to negate EA present verbal sentences is the non-discontinuous independent particle *muš*, as in (21a–b). These two particles do not exhibit any semantic or pragmatic differences, but they differ in their usage frequency. According to Albirini and Binmamoun (2015) and based on the performance of the control participants, the discontinuous particle *ma-š* is more widely used in the EA output, compared to the non-discontinuous particle *muš*. Although the latter was somewhat present in the data produced by the control participants, it was not presented as a negator of the EA verbal present tense in the L2 participants’ textbooks. It was only introduced as a negator of the nominal-verbless present

construction, as in (22a–b). Thus, it was not counted toward the total correct production tokens of the verbal present negation structure, as will be discussed further in Section 3.6.2.4 below.

- |      |    |   |  |  |
|------|----|---|--|--|
| (20) | a. | <i>ʔil-walad</i><br>the-boy<br>“The boy does not study.”                          |  | <i>ma-bi-yi-zākir-š</i><br>Neg-Hb-Imperf.3.s.m-study-Neg |
|      | b. | <i>ʔil-bint</i><br>the-girl<br>“The girl does not study.”                         |  | <i>ma-bi-t-zākir-š</i><br>Neg-Hb-Imperf.3.s.f-study-Neg  |
| (21) | a. | <i>ʔil-walad</i> <i>muš</i><br>the-boy          Neg<br>“The boy does not study.”  |  | <i>bi-yi-zākir</i><br>Hb-Imperf.3.s.m-study              |
|      | b. | <i>ʔil-bint</i> <i>muš</i><br>the-girl          Neg<br>“The girl does not study.” |  | <i>bi-t-zākir</i><br>Hb-Imperf.3.s.f-study               |
| (22) | a. | <i>ʔil-walad</i> <i>muš</i><br>the-boy          Neg<br>“The boy is not hungry.”   |  | <i>gēḫān</i><br>hungry.s.m                               |
|      | b. | <i>ʔil-bint</i> <i>muš</i><br>the-girl          Neg<br>“The girl is not hungry.”  |  | <i>gēḫān-a</i><br>hungry-s.f                             |

### 3.3.3 Subjunctive subordination

Subordination generally refers to two clauses: independent and dependent. Whereas the independent clause functions as the matrix/main clause, the dependent clause is subordinated or embedded within it. In MSA, subordinate clauses are introduced by various complementizers, one of which is termed *ʔan ʔan-nāṣibah* “the subjunctive particle *to*.” According to LeTourneau (2008), it introduces unasserted and nonfactual propositions that indicate a future contingency. It is subcategorized by predicate verbs denoting a vast range of meanings, such as desire (e.g., *yuhibb* “to like”), wish (e.g., *yurīd* “to want”), possibility (e.g., *yumkin* “to be possible”),

necessity (e.g., *yajib* “to be necessary”), command (e.g., *yaʔmur* “to command”), or fear (e.g., *yaxāf* “to fear”).

Unlike English, the subjunctive subordinate clause in MSA and EA requires the verb to be in its finite form—that is, in full agreement with the coindexed antecedent in the main clause. However, the similarity between the MSA and EA stops here. To illustrate, like English, MSA entails spelling out the complementizer *ʔan* “to,” as in (23a–b). In contrast, the complementizer in EA is obligatorily null, resulting in a serial–verb construction, as in (24a–b). Dropping the complementizer in MSA or spelling it out in EA would render the discourse unauthentic and unnatural, as noted by Alhawary (2021) on some of the L2ers’ output samples in Leddy-Cecere’s (2018) study (see Section 2.4: Chapter 2).

- |      |    |                                       |  |                                   |   |
|------|----|---------------------------------------|--|-----------------------------------|---|
| (23) | a. | <i>ʔal-walad(-u)</i><br>the-boy(-Nom) | <i>yu-rīd(-u)</i><br>Imperf.3.s.m-want(-Indic) | <i>ʔan</i><br>to                  | <i>ya-ʔkul(-a)</i><br>Imperf.3.s.m-eat(-Subj) |
|      |    | “The boy wants to eat.”               |  |                                   |   |
|      | b. | <i>ʔal-bint(-u)</i><br>the-girl(-Nom) | <i>tu-rīd(-u)</i><br>Imperf.3.s.f-want(-Indic) | <i>ʔan</i><br>to                  | <i>ta-ʔkul(-a)</i><br>Imperf.3.s.f-eat(-Subj) |
|      |    | “The girl wants to eat.”              |  |                                   |   |
| (24) | a. | <i>ʔil-walad</i><br>the-boy           | <i>ʕāwiz/ʕāyiz</i><br>wanting.s.m              | <i>yā-kul</i><br>Imperf.3.s.m-eat |   |
|      |    | “The boy wants to eat.”               |  |                                   |   |
|      | b. | <i>ʔil-bint</i><br>the-girl           | <i>ʕāwz-a/ʕāyiz-a</i><br>wanting-s.f           | <i>tā-kul</i><br>Imperf.3.s.f-eat |   |
|      |    | “The girl wants to eat.”              |  |                                   |   |

MSA and EA require the subject of the subordinate clause in the subjunctive construction to be null (i.e., big PRO). However, the production data of the control participants revealed that the EA subjunctive subordination construction can be expressed by means of using the construction of *ʔinn* “than,” subcategorized by a nonfactual proposition verb or subject-noun predicate and

followed by a nominal sentence containing a verb predicate. When this structure is used, the subject of the nominal sentence in the subordinate clause has to be overtly spelled out and coindexed with its antecedent in the matrix clause. The subject can be an overt subject pronoun, as in (25a–b), or a subject enclitic, as in (26a–b).

- (25) a. *ʔil-walad*            *ʕāwiz/ʕāyiz*            *ʔinn*    *huwwa*            *yā-kul*  
the-boy                    wanting.s.m            that    he                    Imperf.3.s.m-eat  
“The boy wants that he eats.”
- b. *ʔil-bint*                *ʕāwz-a/ʕāyiz-a*            *ʔinn*    *hiyya*                *tā-kul*  
the-girl                    wanting-s.f            that    she                    Imperf.3.s.f-eat  
“The girl wants that she eats.”
- (26) a. *ʔil-walad*            *ʕāwiz/ʕāyiz*                            *ʔinn-u*                            *yā-kul*  
the-boy                    wanting.s.m                            that-he                            Imperf.3.s.m-eat  
“The boy wants that he eats.”
- b. *ʔil-bint*                *ʕāwz-a/ʕāyiz-a*                            *ʔinna-hā*                            *tā-kul*  
the-girl                    wanting-s.f                            that-she                            Imperf.3.s.f-eat  
“The girl wants that eats.”

Note that the subjunctive subordination construction can be optionally substituted by a verbal noun derived from the subjunctive subordinate verb. This process transforms the complex sentences in (23a–b) into simple sentences, as in (27a–b). Accordingly, any instances of using the verbal noun structure instead of the subjunctive subordination structure were not counted toward the total production tokens of the latter, as will be discussed in Section 3.6.2.5 below.

- (27) a. *ʔal-walad(-u)*                            *yu-rīd(-u)*                            *ʔal-ʔakl(a)*  
the-boy(-Nom)                            Imperf.3.s.m-want(-Indic)                            the-eating.VN(-Acc)  
“The boy wants [to] eat.”
- b. *ʔal-bint(-u)*                            *tu-rīd(-u)*                            *ʔal-ʔakl(-a)*  
the-girl(-Nom)                            Imperf.3.s.f-want(Indic)                            the-eating.VN(-Acc)  
“The girl wants [to] eat.”

### 3.3.4 Cause–effect subordination

Another type of particle subordinating dependent clauses within matrix clauses is the conjunction *liʔanna* and the EA conjunction *ʕašān/ʕalašān* “because.” The particle *liʔanna* consists of the preposition *li* and the complementizer *ʔanna*. This combination provides the meaning of “because.” This conjunction is always followed by either a nominal-verbless sentence or a nominal sentence containing a verb predicate. In both types of the subordinate nominal sentences, the subject must be coindexed with its antecedent in the matrix clause. It may be spelled out as an overt noun, as in (28a–b) and (29a–b), for the nominal-verbless construction and the nominal construction containing a verb predicate, respectively. The subject may also be realized as an enclitic attached to the conjunction *liʔanna* “because,” as in (30a–b) and (31a–b), for the nominal-verbless construction and the nominal construction containing a verb predicate, respectively.

- (28) a. *ʔal-walad(-u) ya-drus(-u) liʔanna ʔal-walad(-a) mujtahid(-un)*  
the-boy(-Nom) Imperf.3.s.m-study(-Indic) because the-boy(-Acc) hardworking(-Nom)  
“The boy studies because the boy is hardworking.”
- b. *ʔal-bint(-u) ta-drus(-u) liʔanna ʔal-bint(-a) mujtahid-at(-un)*  
the-girl(-Nom) Imperf.3.s.f-study(-Indic) because the-girl(-Acc) hardworking-s.f(-Nom)  
“The girl studies because the girl is hardworking.”
- (29) a. *ʔal-walad(-u) najaha liʔanna ʔal-walad(-a) ḏākara*  
the-boy(-Nom) Perf.succeeded.3.s.m because the-boy(-Acc) Perf.studied.3.s.m  
“The boy succeeded because the boy studied.”
- b. *ʔal-bint(-u) najah-at liʔanna ʔal-bint(-a) ḏākar-at*  
the-girl(-Nom) Perf.succeeded-3.s.f because the-girl(-Acc) Perf.studied-3.s.f  
“The girl succeeded because the girl studied.”
- (30) a. *ʔal-walad(-u) ya-drus(-u) liʔanna-hu mujtahid(-un)*  
the-boy(-Nom) Imperf.3.s.m-study(-Indic) because-he hardworking.s.m(-Nom)  
“The boy studies because he is hardworking.”



- b. *ʔal-bint(-u)*     *ta-drus(-u)*     *liʔanna-hā*     *mujtahid-at(-un)*  
the-girl(-Nom) Imperf.3.s.f-study(-Indic) because-she hardworking-s.f(-Nom)  
“The girl studies because she is hardworking.”
- (31) a. *ʔal-walad(-u)*     *najaḥa*     *liʔanna-hu*     *ḍākara*  
the-boy(-Nom) Perf.succeeded.3.s.m because-he Perf.studied.3.s.m  
“The boy succeeded because he studied.”
- b. *ʔal-bint(-u)*     *najaḥ-at*     *liʔanna-hā*     *ḍākar-at*  
the-girl(-Nom) Perf.succeeded-3.s.f because-she Perf.studied-3.s.f  
“The girl succeeded because she studied.”

Regarding the cause–effect relation in EA, it can be conveyed by either the conjunctions *liʔann* or *ʕašān/ʕalašān* “because.” Notwithstanding the phonological differences between MSA and EA, the subordinate clause headed by the conjunction *liʔann* “because” behaves similarly in both varieties. Sentences (28–31) in the MSA variety are restated in (32–35) in the EA variety.

- (32) a. *ʔil-walad*     *bi-yi-zākir*     *liʔann*     *ʔil-walad*     *mugtahid*  
the-boy Hb-Imperf.3.s.m-study because the-boy hardworking.s.m  
“The boy studies because the boy is hardworking.”
- b. *ʔal-bint*     *bi-t-zākir*     *liʔann*     *ʔil-bint*     *mugtahid-a*  
the-girl Hb-Imperf.3.s.f-study because the-girl hardworking-f.s  
“The girl studies because the girl is hardworking.”
- (33) a. *ʔil-walad*     *nigḥ*     *liʔann*     *ʔil-walad*     *zākar*  
the-boy Perf.succeeded.3.s.m because the-boy Perf.studied.3.s.m  
“The boy succeeded because the boy studied.”
- b. *ʔal-bint*     *nigḥ-it*     *liʔann*     *ʔil-bint*     *zakar-it*  
the-girl Perf.succeeded-3.s.f because the-girl Perf.studied-3.s.f  
“The girl succeeded because the girl studied.”
- (34) a. *ʔil-walad*     *bi-yi-zākir*     *liʔann-u*     *mugtahid*  
the-boy Imperf.3.s.m-study because-he hardworking.s.m  
“The boy studies because he is hardworking.”

- |      |    |  |   |                                  |                                       |
|------|----|--|---|----------------------------------|---------------------------------------|
|      | b. | <i>ʔil-bint</i><br>the-girl                    | <i>bi-t-zākir</i><br>Imperf.3.s.f-study | <i>liʔanna-hā</i><br>because-she | <i>mugtahid-a</i><br>hardworking-s.f  |
|      |    | “The girl studies because she is hardworking.” |   |                                  |                                       |
| (35) | a. | <i>ʔil-walad</i><br>the-boy                    | <i>nigih</i><br>Perf.succeeded.3.s.m    | <i>liʔann-u</i><br>because-he    | <i>zākar</i><br>Perf.studied.3.s.m    |
|      |    | “The boy succeeded because he studied.”        |   |                                  |                                       |
|      | b. | <i>ʔil-bint</i><br>the-girl                    | <i>nigh-it</i><br>Perf.succeeded-3.s.f  | <i>liʔanna-hā</i><br>because-she | <i>zakar-it</i><br>Perf.studied-3.s.f |
|      |    | “The girl succeeded because she studied.”      |   |                                  |                                       |

However, when the EA conjunction *ʔašān/ʔalašān* “because” is used to convey the cause–effect relation, the subordinate clause can also be followed by a verbal sentence, as in (36a–b).

- |      |    |   |   |   |                                       |
|------|----|---|---|---|---------------------------------------|
| (36) | a. | <i>ʔil-walad</i><br>the-boy             | <i>bi-yi-zākir</i><br>Hb-Imperf.3.s.m-study | <i>ʔašān/ʔalašān</i><br>because/in order to | <i>yi-ngh</i><br>Imperf.3.s.m-succeed |
|      |    | “The boy studies in order to succeed.”  |   |   |                                       |
|      | b. | <i>ʔil-bint</i><br>the-girl             | <i>bi-t-zākir</i><br>Hb-Imperf.3.s.f-study  | <i>ʔašān/ʔalašān</i><br>because/in order to | <i>ti-ngh</i><br>Imperf.3.s.f-succeed |
|      |    | “The girl studies in order to succeed.” |   |   |                                       |

It should be noted that the cause–effect relation can be optionally conveyed by the phrase *bisabab* “because of” in MSA and the conjunction *ʔašān/ʔalašān* “because of” in EA. When they are used to convey the meaning of “because of,” they behave similarly, and the construction in which they occur is considered simple rather than complex. This is so because these two phrases are considered the first part of the *ʔiḍāfah* construction and are always followed by a noun, as in (37a–b) and (38a–b) for MSA and EA, respectively. Accordingly, any production tokens of a conjunction that conveys the meaning of “because of” and followed by an *ʔiḍāfah* phrase were not counted toward the total production tokens of the cause–effect subordination structure, as will be discussed further in Section 3.6.2.6 below.

- |      |    |  |   |  |                                |
|------|----|--|---|--|--------------------------------|
| (37) | a. | <i>ʔal-walad(-u)</i><br>the-boy(-Nom)      | <i>qaliq(-un)</i><br>nervous.s.m(-Nom)    | <i>bisabab(-i)</i><br>because of(-Gen) | <i>ʔal-ixtibār</i><br>the-exam |
|      |    | “The boy is nervous because of the exam.”  |   |  |                                |
|      | b. | <i>ʔal-bint(-u)</i><br>the-girl(-Nom)      | <i>qaliq-at(-un)</i><br>nervous-s.f(-Nom) | <i>bisabab(-i)</i><br>because of(-Gen) | <i>ʔal-ixtibār</i><br>the-exam |
|      |    | “The girl is nervous because of the exam.” |   |  |                                |
| (38) | a. | <i>ʔil-walad</i><br>the-boy                | <i>ʔalʔān</i><br>nervous.s.m              | <i>ʕašān/ʕalašān</i><br>because of     | <i>ʔil-imiṭḥān</i><br>the-exam |
|      |    | “The boy is nervous because of the exam.”  |   |  |                                |
|      | b. | <i>ʔil-bint</i><br>the-girl                | <i>ʔalʔān-a</i><br>nervous-s.f            | <i>ʕašān/ʕalašān</i><br>because of     | <i>ʔil-imiṭḥān</i><br>the-exam |
|      |    | “The girl is nervous because of the exam.” |   |  |                                |

### 3.4 Materials

As discussed in Chapter 2, one of the major limitations of Arabic L2 studies that examined the effectiveness of the IA pertains to their reliance on previously prepared skits and presentations performed by Arabic L2ers. The current study avoided this limitation by eliciting spontaneous production data of the target structures from all participants. Thus, narrative and manipulation tasks were designed and conducted in elicitation sessions that were held in the form of in-person one-on-one structured interviews with all participants. Before discussing the study tasks, it is crucial to provide a comprehensive analysis of the timing and frequency of the target structures in the L2 participants’ instructional textbooks.

#### 3.4.1 Instructional input of the target structures

This section provides a comprehensive analysis of the timing and frequency of the target structures in the instructional textbook series used by the IPG and UIPG—that is, Part I and Part II of the third edition of *Al-Kitāb fī Ta‘allum Al-‘Arabiyya* (Brustad et al., 2011, 2013) and Part I and Part II of the second edition of *Al-Kitāb fī Ta‘allum Al-‘Arabiyya* (Brustad et al., 2004,

2007). It is crucial to stress the inherent complexity involved in isolating and controlling for all aural input received by the participants, within the controlled environment of the classroom and in the unpredictable external surroundings. As a consequence, this analysis meticulously directs its exclusive focus toward the visual input received by the participants through their instructional textbooks.

#### ***3.4.1.1 Instructional input of the past tense***

Since MSA and EA converge in the past tense structure, its presentation timing and frequency in the input were analyzed regardless of the variety in which it was presented. Table 3.5 displays when and how often the past tense was presented in both instructional textbook series used by the IPG and UIPG. The past tense, with both forms of 3.s.m and 3.s.f, was formally introduced at the same time in both series (in Lesson 8) and was sufficiently recycled thereafter in the second edition. However, in the third edition, it was not introduced at all in Lesson 9 and was introduced with the 3.s.m form only in Lesson 10.

**Table 3.5** Timing and Frequency of the Past Tense in the Instructional Textbooks used by the IPG and UIPG

Unit/Lesson	3.s.m	3.s.f	3.s.m	3.s.f
1				
2		x		
3				^
4			^	
5	^	^		^
6	^			
7	^	^	x	
8	√	√	√	√
9			x	x
10	x		x	x
11	x	x	x	x
12	x	x	x	x
13	x	x	x	x
14	x	x	x	x
15	x	x	x	x
16	x	x	x	x
17	x	x	x	x
18	x	x	x	x
19	x	x	x	x
20	x	x	x	x
21	x	x	x	x
22	x	x	x	x
23	x	x	x	x
24	*	*	x	x
25	*	*	x	x
26	*	*	x	x
27	*	*	x	x
28	*	*	x	x
29	*	*	x	x
30	*	*	x	x

*Note*<sup>1</sup>. The shaded left half of the table is for Part I and Part II of the third edition of *Al-Kitāb fī Ta'allum Al-'Arabiyya* (Brustad et al., 2011, 2013). The unshaded right half of the table is for Part I and Part II of the second edition of *Al-Kitāb fī Ta'allum Al-'Arabiyya* (Brustad et al., 2004, 2007).

*Note*<sup>2</sup>. √ = structure is the focus of instruction; x = structure is not the focus of instruction but occurs in the lesson 4 or more times; ^ = structure is not the focus of instruction and occurs less than 4 times, \* = no lesson (following Alhawary, 2009a, 2019).

### ***3.4.1.2 Instructional input of the present tense***

Table 3.6 displays when and how often the present tense structure was presented in both instructional textbook series used by the IPG and UIPG participants. The MSA present tense, with both forms of 3.s.m and 3.s.f, was formally introduced at the same time in both textbook series (in Lesson 4) and was mostly maintained thereafter—although the second edition has a slightly more frequent presentation of this structure in the first three lessons. However, the EA aspectual marker of present habituality was not frequently recycled in third edition—although it was formally introduced as early as Lesson 4.

**Table 3.6** Timing and Frequency of the Present Tense in the Instructional Textbooks Used by the IPG and UIPG

Unit/Lesson	MSA		EA		MSA	
	3.s.m	3.s.f	3.s.m	3.s.f	3.s.m	3.s.f
1	^	x	^	^	x	x
2	^	x			x	x
3	^	x	^		x	X
4	√	√	√	√	√	√
5	x	x	^	^	x	x
6	x	x	^		x	x
7	x	x	^		x	x
8	x	^			x	x
9	√	^	^	^	x	x
10	x	x	^		x	x
11	x	x	^		x	x
12	x	x	^	^	x	^
13	x	^	^	^	x	x
14	x	x	x	^	x	x
15	x	x	^	^	x	x
16	x	x	^		x	x
17	x	x			x	x
18	x	x	^	^	x	x
19	x	x	^	^	x	x
20	x	x			x	x
21	x	x	^	^	x	x
22	x	x			x	x
23	x	x	^		x	x
24	*	*	*	*	x	x
25	*	*	*	*	x	x
26	*	*	*	*	x	x
27	*	*	*	*	x	x
28	*	*	*	*	x	x
29	*	*	*	*	x	x
30	*	*	*	*	x	x

*Note*<sup>1</sup>. The shaded left half of the table is for Part I and Part II of the third edition of *Al-Kitāb fī Ta'allum Al-'Arabiyya* (Brustad et al., 2011, 2013). The unshaded right half of the table is for Part I and Part II of the second edition of *Al-Kitāb fī Ta'allum Al-'Arabiyya* (Brustad et al., 2004, 2007).

*Note*<sup>2</sup>. √ = structure is the focus of instruction; x = structure is not the focus of instruction but occurs in the lesson 4 or more times; ^ = structure is not the focus of instruction and occurs less than 4 times, \* = no lesson (following Alhawary, 2009a, 2019).

### ***3.4.1.3 Instructional input of the past negation***

Table 3.7 displays the when and how often the past negation was introduced in both instructional textbook series used by the IPG and UIPG. Notwithstanding the high frequency of the MSA negation particle *lam* “did not”, it was formally introduced as late as Lesson 13 in both textbook series and was mostly maintained thereafter. Conversely, the MSA negation particle *mā* “did not,” which is more marked than the former, was introduced at an earlier stage, particularly in Lesson 8. Despite its early introduction and its markedness, it was insufficiently recycled afterwards. Similarly, the EA past negation construction *ma-perfect-š* “did not” was implicitly introduced as early as Lesson 8 in the third edition and was not sufficiently recycled thereafter.



**Table 3.7** Timing and Frequency of the Past Negation in the Instructional Textbooks Used by the IPG and UIPG

Unit/Lesson	MSA		EA	MSA	
	<i>lam</i>	<i>mā</i>	<i>ma-perfect-š</i>	<i>lam</i>	<i>mā</i>
1					
2	^				
3					
4				^	
5					
6					
7					^
8		√	^		√
9					^
10		^	^		
11				^	^
12	^	^	^	^	
13	√	^	^	√	^
14	x	^	^	^	
15	x	^	^	x	
16	x	^		^	^
17	x	^		x	^
18	x		^	x	
19	x	^		^	
20	x	^		x	
21	x	^	^	x	
22	x	^	x	^	^
23	x	^		x	
24	*	*	*	x	
25	*	*	*	x	
26	*	*	*	x	
27	*	*	*	x	
28	*	*	*	x	
29	*	*	*	x	
30	*	*	*	x	^

*Note*<sup>1</sup>. The shaded left half of the table is for Part I and Part II of the third edition of *Al-Kitāb fī Taʿallum Al-ʿArabiyya* (Brustad et al., 2011, 2013). The unshaded right half of the table is for Part I and Part II of the second edition of *Al-Kitāb fī Taʿallum Al-ʿArabiyya* (Brustad et al., 2004, 2007).

*Note*<sup>2</sup>. √ = structure is the focus of instruction; x = structure is not the focus of instruction but occurs in the lesson 4 or more times; ^ = structure is not the focus of instruction and occurs less than 4 times, \* = no lesson (following Alhawary, 2009a, 2019).

#### ***3.4.1.4 Instructional input of the verbal present negation***

Table 3.8 displays when and how often the verbal present negation was introduced in both instructional textbook series used by the IPG and UIPG. In both textbook series, the MSA negation particle *lā* “does/do not” was formally introduced in Lesson 4 and was maintained thereafter. As for the EA verbal present negation construction *ma-bi-imperfect-š* “does/do not,” it was not explicitly introduced in any lesson of the third edition and was not sufficiently recycled throughout.

**Table 3.8** Timing and Frequency of the Verbal Present Negation in the Instructional Textbooks Used by the IPG and UIPG

Unit/Lesson	MSA	EA	MSA
	<i>lā</i>	<i>ma-bi-imperfect-š</i>	<i>lā</i>
1	^	^	^
2	^		
3			^
4	√	^	√
5	x	^	x
6	x	^	x
7	^		^
8	x	^	^
9	x		x
10	x		x
11	x		^
12	x		
13	x		x
14	x	^	x
15	x	^	x
16	x		^
17	x		x
18	x		x
19	x		x
20	x		^
21	x	^	^
22	x	^	x
23	x		x
24	*	*	x
25	*	*	x
26	*	*	x
27	*	*	x
28	*	*	x
29	*	*	x
30	*	*	x

*Note*<sup>1</sup>. The shaded left half of the table is for Part I and Part II of the third edition of *Al-Kitāb fī Ta'allum Al-'Arabiyya* (Brustad et al., 2011, 2013). The unshaded right half of the table is for Part I and Part II of the second edition of *Al-Kitāb fī Ta'allum Al-'Arabiyya* (Brustad et al., 2004, 2007).

*Note*<sup>2</sup>. √ = structure is the focus of instruction; x = structure is not the focus of instruction but occurs in the lesson 4 or more times; ^ = structure is not the focus of instruction and occurs less than 4 times, \* = no lesson (following Alhawary, 2009a, 2019).

#### ***3.4.1.5 Instructional input of the subjunctive subordination***

Table 3.9 displays when and how often the MSA and EA subjunctive subordination structures were introduced in the instructional textbook series used by the IPG and UIPG. In both editions, the MSA construction was formally introduced in Lesson 10 and was sufficiently recycled thereafter. The EA construction was formally introduced in Lessons 4 and 10 of the third edition, but it was not sufficiently recycled throughout.

**Table 3.9** Timing and Frequency of the Subjunctive Subordination in the Instructional Textbooks Used by the IPG and UIPG

Unit/Lesson	MSA	EA	MSA
	ʔan + imperfect	verb + imperfect	ʔan + imperfect
1			
2			
3			
4	^	√	
5		^	^
6	^	^	
7	^		
8	x	x	x
9	^	^	^
10	√	√	√
11	x		x
12	x	x	x
13	x	x	x
14	x	^	x
15	^	^	x
16	x	^	x
17	x	^	x
18	x	^	x
19	x	^	x
20	x	^	x
21	x	^	x
22	x	^	x
23	x	^	x
24	*	*	x
25	*	*	x
26	*	*	x
27	*	*	x
28	*	*	x
29	*	*	x
30	*	*	x

*Note*<sup>1</sup>. The shaded left half of the table is for Part I and Part II of the third edition of *Al-Kitāb fī Taʿallum Al-ʿArabiyya* (Brustad et al., 2011, 2013). The unshaded right half of the table is for Part I and Part II of the second edition of *Al-Kitāb fī Taʿallum Al-ʿArabiyya* (Brustad et al., 2004, 2007).

*Note*<sup>2</sup>. √ = structure is the focus of instruction; x = structure is not the focus of instruction but occurs in the lesson 4 or more times; ^ = structure is not the focus of instruction and occurs less than 4 times, \* = no lesson (following Alhawary, 2009a, 2019).

### **3.4.1.6 Instructional input of the cause–effect subordination**

Table 3.10 displays when and how often the cause–effect subordination with the conjunction *li?anna* “because” and the EA conjunction *ʒašān/ʒalašān* “because” was introduced in the instructional textbook series used by the IPG and UIPG. In both editions, the conjunction *li?anna* “because” was formally introduced in Lesson 6 and was sufficiently recycled subsequently. However, the EA conjunction was neither formally introduced nor sufficiently recycled in the third edition.

**Table 3.10** Timing and Frequency of the Cause–Effect Subordination with the Conjunction *liʔanna* “because” and the EA Conjunction *ʕašān/ʕalašān* “because” in the Instructional Textbooks Used by the IPG and UIPG

Unit/Lesson	MSA	EA	MSA
	<i>liʔanna</i>	<i>ʕašān/ʕalašān</i>	<i>liʔanna</i>
1			
2			
3			
4			^
5			
6	√	^	√
7		^	x
8	^	x	x
9	^	^	x
10	x	^	
11	x	^	^
12	^	^	x
13	x		^
14	x		^
15	x		x
16	x		x
17	x		x
18	x	^	x
19	x		^
20	x	^	^
21	x		^
22	x		^
23	x		^
24	*	*	x
25	*	*	x
26	*	*	x
27	*	*	x
28	*	*	x
29	*	*	x
30	*	*	x

*Note*<sup>1</sup>. The shaded left half of the table is for Part I and Part II of the third edition of *Al-Kitāb fī Taʕallum Al-ʕArabiyya* (Brustad et al., 2011, 2013). The unshaded right half of the table is for Part I and Part II of the second edition of *Al-Kitāb fī Taʕallum Al-ʕArabiyya* (Brustad et al., 2004, 2007).

*Note*<sup>2</sup>. √ = structure is the focus of instruction; x = structure is not the focus of instruction but occurs in the lesson 4 or more times; ^ = structure is not the focus of instruction and occurs less than 4 times, \* = no lesson (following Alhawary, 2009a, 2019).

### 3.4.2 Tasks

Two narrative tasks and one manipulation task were designed to elicit spontaneous production data from the participants. Révész (2009) laid out the benefit of using pictures to elicit spontaneous output by emphasizing that tokens of target structures were ample and sufficient in the data obtained from L2ers who performed picture-based tasks, compared to their counterparts who performed tasks without pictures.

Various criteria were carefully considered when designing the three study tasks. First, the drawings of the narrative tasks were connected and contained detailed actions, promoting long sketches of oral production data. Second, because unclear or inconspicuous pictures would increase the cognitive load on participants who would spend more time figuring out and linking the actions in the pictures, a professional cartoonist was hired to render all pictures conspicuous and clear. Third, to eliminate the possibility of confounding the data, the pictures were designed to elicit the target structures in obligatory contexts only. Fourth, to not promote conscious thinking of the target structures in the elicitation sessions, some distractor pictures unrelated to the target structures were included, particularly in the manipulation task. Fifth, to control for the variable pertaining to the levels of familiarity and difficulty of the potential words that should be used when performing the tasks, all designed pictures entailed using basic words that were mainly introduced in Part I of the instructional textbooks used by both L2 groups. Sixth, the tasks were designed in a cautious and neutral way that was neither sensitive nor offensive to anyone—that is, they did not involve any stereotypical bias against any gender, race, religion, or culture. Finally, to ensure maximum comprehension of the tasks, the task instructions were provided in English and presented in a generally short format. Sections 3.4.2.1 and 3.4.2.2 provide a description of the narrative and manipulation tasks, respectively.



### ***3.4.2.1 Narrative tasks***

A narrative task is typically defined as the telling of a story based on various stimuli, the most typical of which are picture sequences. In the current study, two narrative tasks were developed to measure the fluency and accuracy constructs in the production of the past tense and its negation, as well as the present tense and its verbal negation, in the 3.s.m and 3.s.f contexts.

To elicit the target structures of the past tense and its negation, the participants were provided with two separate 13-day calendars containing activities that were carried out by a male character (see Figure 1c: Appendix C) and a female character (see Figure 2c: Appendix C) during their vacations that took place in a previous month. The participants were instructed (in English) to narrate day by day what the male and female characters did and what they did not do during their 13-day vacations. A distractor was integrated into the task instruction, where the participants were asked to tell which one of the vacations they liked the most and why.

To elicit the target structures of the present tense and its verbal negation, the participants were provided with two separate daily routine tables containing activities carried out by a male character (see Figure 1d: Appendix D) and a female character (see Figure 2d: Appendix D). The participants were instructed (in English) to narrate what the male and female characters do and what they do not do at specific times of the day. A distractor was integrated into the task instruction, where the participants were asked to tell which one of the two daily routines resembled theirs and which one seemed more interesting and why.

### ***3.4.2.2 Manipulation task***

In SLA settings, a manipulation task typically refers to a situation in which the oral output of L2ers is predictable because their productions (i.e., responses) are controlled by questions (i.e.,

stimuli). In the current study, one manipulation task was designed to examine the complexity construct in the production of the subjunctive and cause–effect subordinations.

The participants were provided with 15 sets of pictures, each of which had two connected or unconnected pictures. They were asked a specific question for each set, and they were asked to provide a full Arabic sentence when describing each set. As a distractor, the pictures of the two target subordination structures were collapsed together and presented to the participants as one task. Another distractor was implemented by integrating five nontarget structure sets of pictures into all picture sets (see Appendix E for all target and nontarget sets of pictures).

### **3.5 Data Collection**

Various steps were followed in the process of data collection. Upon obtaining an exemption from the University of Michigan’s Institutional Review Board (IRB) to conduct the study, the study materials were virtually piloted (via Zoom) on four Arabic L2ers in an integrated Arabic L2 program at a U.S. university. This step was critical because it yielded valuable information related to the clarity of the tasks, the suitability of their instructions, and the recruitment procedure of L2 participants. The pilot study revealed several flaws that were resolved accordingly. The first limitation pertained to the clarity of the hand-drawn task pictures. Although they fulfilled their purpose in terms of eliciting enough tokens of the target structures in obligatory contexts, the inconspicuous and unclear nature of the drawings was problematic, as pointed out by the pilot study participants. Accordingly, a professional cartoonist was hired to redraw all the pictures and render them clearer. The second flaw revolved around the difficulty and time consumption in the virtual recruitment of the L2 participants. Consequently, a decision was made to personally travel to the two recruitment sites and collect the data for the actual study in the form of in-person interviews with all L2 participants.

The second step entailed contacting the administrators/directors of the two L2 recruitment sites (i.e., IP and UIP) to gather more information about their Arabic programs (see Section 3.2.1 above) and to obtain their approval to conduct the study on the L2ers enrolled in their programs. Upon obtaining the permission from the directors, two visit rounds were paid to each one of the two recruitment sites. The first visit served to collect data from the second-year and third-year participants toward the end of the first semester of the 2021–2022 academic year. As for the first-year participants, it was necessary to collect data from them after they had sufficient exposure to all target structures in their instructional textbooks. Therefore, a second visit was paid to the two L2 recruitment sites to collect data from the first-year participants toward the end of the second semester of the 2021–2022 academic year.

Before visiting the two recruitment sites, an invitation email containing brief information about the study and a link to the LHQ was sent to the two Arabic program coordinators, who subsequently forwarded it to the L2ers enrolled in their programs. However, this step was unsuccessful because only a small number of L2ers filled out the online LHQ. Consequently, arrangements were made with the coordinators to visit the Arabic classes in their programs in-person. At each class visit, the study was briefly introduced, without any mention of the underlying design of the study. Rather, a general explanation of the tasks was provided (e.g., you will be describing various pictures). To promote participation, the L2ers were assured their participation was completely voluntary, involved no risk of any kind at all, would provide them with an opportunity to practice their Arabic skills with a native speaker, and they would receive \$20, as a token of appreciation and compensation for their time and participation in the study.

Then, the consent form, the LHQ, and the sign-up sheet for the one-on-one meetings were distributed to those who showed interest in participation. Subsequently, the LHQ forms

were collected. Among all potentially volunteered participants, confirmation emails/texts of the time and location of the one-on-one meetings were sent to those who met the primary inclusion criteria for L2 participant recruitment: those being English L1 speakers and nonheritage speakers of Arabic, as well as those who had not studied Arabic in previous programs for more than 6 months and had not resided in any Arabic-speaking countries for more than 3 months.

To control for the noise-related variable, which could have possibly distracted the participants and affected the quality of recordings, all one-on-one meetings took place in private and quiet offices, which were conveniently commuted to by all the participants at both recruitment sites. During the meetings, the participants were asked to take the first 20 minutes to finish the reading proficiency-placement test. Then, they received a brief description of the tasks they would perform and were asked to speak as much as they could in Arabic, without specifying the variety they should use. Although some IPG participants asked whether they should use MSA or EA, they were instructed to speak in either variety they preferred since both varieties belong to the same language.

The three study tasks were presented to the participants in the following sequence: the narrative task of the past tense and its negation was presented toward the beginning of the interview, the manipulation task of subjunctive and cause–effect subordinations was presented toward the middle of the interview, and the narrative task of the present tense and its negation was presented toward the end of the interview.

To preserve the production data from any bias or contamination that could result from errors in the data collection procedures, the elicitation instruction of each task (see section 3.4.2 above) was held consistently across all participants. Moreover, to promote continuous and fluent production in the narrative tasks, no interference from the interviewer took place—even if the

participants could not produce certain Arabic words. This procedure was essential because interactional and interventional instances would have affected the participants' thoughts and sent some signals to the participants to prioritize one of the proficiency constructs at the expense of others (e.g., fluency over accuracy, accuracy over fluency). Nonetheless, indirect interference was necessary, particularly when some participants employed the avoidance strategy to skip the target structures. Still, this kind of interference was minimal to keep the participants unaware of the target structures and to prevent any hindrance to their fluency. In the manipulation task, however, a large number of participants could not recall or were not familiar with certain verbs, adjectives, and nouns that were obligatory in contexts. To promote production of full Arabic complex utterances, those participants were somewhat helped with some content words, particularly nouns and adjectives. However, none of the relevant parts of the target structures (e.g., verbs, conjunctions, particles) were provided.

Regarding the data collection from the control participants, the one-on-one meetings with them took place at convenient venues for them toward the end of the second term of the 2021–2022 academic year. The above-mentioned data collection procedure that was carried out with the L2 participants was followed with the control participants as well. An exception was that the reading proficiency-placement test was not administered to them because their proficiency level was determined as native by default.

To keep the participants' identities confidential, real names and other identifying information have remained and will remain anonymous. Additionally, all data obtained from all participants were preserved safely by means of keeping them on password-protected and encrypted files on a personal computer, whose access has been and will be limited to the researcher and advisor only.

## 3.6 Data Coding

### 3.6.1 General coding criteria

All elicitation sessions were audio-recorded. A total of approximately 45 hours of audio recordings constituted the production data for the current study. The data were transcribed and coded using MAXQDA software. Some data were not transcribed, including hesitations, repetitions, responses to distractors, and self-corrections (except for last attempts). Each data sample was checked several times to ensure accuracy of the transcription and coding processes. Examination of all data was carried out, regardless of whether verbs were produced in nominal-verbal sentences (i.e., SVO word order) or verbal sentences (i.e., VSO word order), regardless of whether the subjects were overt or null, and regardless of the correct/incorrect production or lack of suffixes for mood (on present tense forms) and case (on nouns and adjectives).

Certain types of mistakes were tolerated—and hence were not coded as incorrect tokens. Those mistakes include metathesis (e.g., \**yaḏbah* instead of *yaḏhab* “to go,” \**yaʃfal* instead of *yafʃal* “to do”), wrong pronunciation of lexemes whose meanings were easily retrieved from the context (e.g., \**yaḏxun* / \**yudaxxil* instead of *yudaxxin* “to smoke”), incorrect usage of prepositions (e.g., \**ya-staxdim bi ʔal-ḥammām* instead of *ya-staxdim ʔal-ḥammām* “he uses the bathroom”), and incorrect usage of the definite article (e.g., \**fi ʔal-yawm ʔal-xamīs* instead of *fi yawm ʔal-xamīs* “on Thursday”). If the surface pronoun/lexical subject was produced incorrectly, but the verb agrees correctly with the referent subject in context, then the production was coded as partially incorrect (e.g., \**huwa ḏahab-at* “he went” instead of *hiya ḏahab-at* “she went”).

The complexity construct can provide a good indication of the extent to which L2ers can expand a single sentence while maintaining sustainability of grammatically error-free production. Accordingly, any single morphosyntactic error (e.g., tense and agreement errors) in

the matrix or subordinate clauses of the subjunctive or cause–effect subordinations renders the whole sentence ungrammatical and hence completely incorrect.

One important variable in the current study revolves around the authenticity of every data token, whether it should be treated as an MSA or EA sentence. Recall, both varieties exhibit many similarities across all language domains. However, they diverge in almost all the six target structures under study. Therefore, to control for the authenticity variable, specifically detailed measures were developed for each target structure. It is, however, crucial to stress the point that the phonological differences between MSA and EA were excluded from the authenticity judgment (e.g., *jiddan* vs. *giddan* “very,” *liʔanna-hu* vs. *liʔann-ū* “because he”). This is so because the current study is a morphosyntactic study in nature, and those phonological differences do not have any impact on the target structures. The developed measures for authenticity focus mainly on the verbs produced. Based on the textbook of *Alif bāʾ* (Brustad et al., 2011), which was used by the IPG and UIPG at the first-year level, MSA and CA diverge in certain verbs, such as the verbs *to go* and *to see/watch*. While the verbs *yīšūf* “to see/watch” and *yirūh* “to go” are exclusively used in CA (including EA), the verbs *yušāhid* “to see/watch” and *yaḏhab* “to go” are exclusively used in MSA. Thus, any productions of the CA verbs in MSA-like sentences, or vice versa, render the whole production unnatural and unauthentic, as will be illustrated below.

To qualitatively and quantitatively analyze how the participants produced the target structures in MSA and EA, the production tokens of every structure were coded under different coding sets, depending on the contextual variety in which the structure was produced (e.g., correct in MSA, correct in EA, partially incorrect in MSA, partially incorrect in EA). Due to the

fact that both MSA and EA converge in the past tense structure, production tokens of this structure were coded without specifying the varieties in which they were produced.

### 3.6.2 Detailed coding criteria

#### 3.6.2.1 Coding criteria for the past tense data

Each data sample was expected to contain at least 16 obligatory-in-context tokens of the past tense structure: eight in the 3.s.m context and eight in the 3.s.f context. Yet, tokens produced to provide extra-descriptions of the pictures were counted toward the overall production of the past tense. Additionally, lack of suppliance of Arabic verbs or production of English verbs instead of their Arabic equivalents were coded as completely incorrect and thus counted toward the total production of the past tense. Furthermore, production tokens of past tense verbs inflected for the first-person singular agreement form (1.s; e.g., *ḍahab-tu* “I went”), second-person singular masculine form (2.s.m; e.g., *ḍahab-ta* “you went”), or second-person singular feminine form (2.s.f; e.g., *ḍahab-ti* “you went”) instead of the 3.s.f agreement form (e.g., *ḍahab-at* “she went”) were coded as completely incorrect.

Because the past progressive aspect/tense was present in the data of some L2 and control participants, production tokens of this aspect/tense were tolerated—and thus counted toward the total production of the past tense. However, any single error in the verb inflectional features for agreement or tense in any part of the past progressive structure renders the produced phrase/sentence ungrammatical and hence coded as completely incorrect (e.g., *\*hiya kānat ya-šrab qahwa* instead of *hiya kānat ta-šrab qahwa* “she was drinking coffee”).

As discussed in Section 3.3.1.1, the EA past progressive is conveyed through the combination of the auxiliary perfect verb/modal *kāna* “was” and the main imperfect verb, which must be preceded by the progressive aspectual marker *bi-* (e.g., *huwwa kān bi-yi-lṣab* “he was



playing”). Accordingly, if the main verb sounds as an EA verb (e.g., *yišūf* “to see/watch” and *yirūh* “to go”), then the aspectual marker *bi-* was expected to be obligatorily produced with the main verb (Brustad et al., 2011). If not, then the token was coded as partially incorrect even if the agreement and tense features were inflected correctly on the verbs (e.g., \**huwwa kān yi-šūf ?il-film* instead of *huwwa kān bi-yi-šūf ?il-film* “he was watching the movie”).

### 3.6.2.2 Coding criteria for the present tense data

Each data sample was expected to contain at least 15 obligatory-in-context tokens of the present tense: seven in the 3.s.m context and eight in the 3.s.f context. Nevertheless, any tokens produced to provide extra-descriptions of the pictures were counted toward the overall production of the present tense. Lack of suppliance of Arabic verbs or production of English verbs instead of their Arabic equivalents were coded as completely incorrect. If a produced present verb was inflected incorrectly for either the wrong agreement or tense, then the whole token was coded as completely incorrect.

To control for the authenticity variable, production tokens of EA-like verbs (e.g., *yišūf* “to see/watch” and *yirūh* “to go”) in MSA-like sentences were coded as partially incorrect (e.g., *huwa ya-ḏhab ?ilā ?al-bayt wa yi-šūf ?at-tilīfīzyūn* instead of *huwa ya-ḏhab ?ilā ?al-bayt wa yu-šāhid ?at-tilīfīzyūn* “he goes home and watches the TV”). Similarly, any productions of MSA-like verbs (e.g., *yušāhid* “to see/watch” and *yaḏhab* “to go”) in EA-like sentences were coded as partially incorrect (e.g., *huwwa bi-yirūh ?il-bēt wa yu-šāhid ?it-tilīfīzyūn* instead of *huwa bi-yirūh ?el-bēt wi yi-šūf ?it-tilīfīzyūn* “he goes home and watches the TV”). Recall, the prefix *bi-* that conveys the habituality aspect is obligatorily used with imperfect verbs in EA. However, the prefix is optionally used in subordinate and conjoined clauses. Accordingly, production tokens of the present tense in EA-like sentences, without the habitual prefix *bi-* in its obligatorily-used

context, were considered unauthentic and hence coded as partially incorrect (e.g., \**hiyya ti-rūh li ʔin-nādī* instead of *hiyya bi-t-rūh li ʔin-nādī* “she goes to the gym”).

### 3.6.2.3 Coding criteria for the past negation data

Each data sample was expected to contain at least 10 obligatory-in-contexts tokens of the past negation structure: five in the 3.s.m context and five in the 3.s.f context. Nevertheless, any tokens produced to provide extra-descriptions of the pictures were counted toward the whole production of the past tense negation structure.

To determine the authenticity of the past negation tokens, whether they were in MSA or EA, the produced verbs were considered first. If the negated verbs are mutually used in both varieties (e.g., *nām* “slept,” *sabaḥ* “swam,” *ʔakal* “ate,” *šarib* “drank”), then the production tokens were treated as either MSA or EA sentences, depending on the negators produced with the verbs. However, if the verbs are exclusively used in either one of the two varieties, then the negator should be the one that is exclusively used in that variety. To illustrate, if either of the two EA-like verbs *šāf* “saw/watched” and *rāḥ* “went” was produced in a negation construction, then the discontinuous EA negator *ma-š* “did not” was expected to be used with the verb produced. Productions of MSA negators *mā* or *lam* with these two verbs render the whole construction unauthentic and hence partially incorrect (e.g., \**huwa mā šāf* instead of *huwa mā šāhada* “he did not watch”). The same rule applies to verbs that are exclusively used in MSA (e.g., *šāhad* “saw,” *ḍahaba* “went”)—that is, production of the EA negator with these two verbs renders the whole construction unauthentic and hence partially incorrect (e.g., \**huwa ma-ḍahab-š* instead of *huwwa ma-raḥ-š* “he did not go”).

Production tokens of the past negation were coded as completely incorrect if they involved the following: correct usage of past negators with verbs inflected for both the wrong

tense and agreement (e.g., \**hiya mā ya-ḏhab* instead of *hiya mā ḏahab-at* “she did not go”); unauthentic negators with verbs inflected for either the wrong agreement or tense (e.g., \**hiya lam yi-rūḥ* instead of *hiya ma-rāḥi-t-š* “she did not go”); correct usage of negators with lack of Arabic verb suppliance (e.g., English verbs instead of Arabic verbs); correct and authentic usage of negators with verbs inflected for the wrong agreement (e.g., \**hiya lam ya-ḏhab* instead of *hiya lam ta-ḏhab* “she did not go”); correct and authentic usage of negators with verbs inflected for the wrong tense (e.g., \**huwa mā ya-ḏhab* instead of *huwa mā ḏahaba* “he did not go”).

#### **3.6.2.4 Coding criteria for the verbal present negation data**

Each data sample was expected to contain at least nine obligatory-in-context tokens of the verbal present negation: five in the 3.s.m context and four in the 3.s.f context. Nevertheless, any tokens produced to provide extra-descriptions of the pictures were counted toward the total production of the verbal present negation structure.

To determine the authenticity of the production tokens of the verbal present negation structure, whether they were in MSA or EA, the produced verbs were considered first. If the negated verbs are mutually used in both varieties (e.g., *yanām* “to sleep,” *yasbaḥ* “to swim,” *yaʔkul* “to eat,” *yašrab* “to drink”), then the production tokens were treated as either MSA or EA sentences, depending on the negators produced with the verbs. In case the EA discontinuous negator *ma-š* “does/do not” was used, then the habitual prefix *bi-* was expected to be produced in its obligatory-used context (see Section 3.3.1.2). If the habitual marker was not used, then the tokens were coded as incorrect. However, if the verbs are exclusively used in either one of the two varieties, then the negator should be the one that is exclusively used in that variety. To illustrate, if either of the two EA-like verbs *yišūf* “to see/to watch” and *yirūḥ* “to go” was produced, then the discontinuous EA negator *ma-š* “does/do not” was expected to be used with

the EA-like verb preceded by the habitual marker *bi-*. Any productions of the MSA negator *lā* “does/do not” with these two verbs render the whole construction unauthentic and hence partially incorrect (e.g., \**huwa lā yi-šūf* instead of *huwa lā yu-šāhid* “he does not watch”). The same rule applies to verbs that are exclusively used in MSA (e.g., *yušāhid* “to watch,” *yaḏhab* “to go”)—that is, any productions of the discontinuous EA negator *ma-š* “does/do not” with these two verbs render the whole construction unauthentic and hence partially incorrect (e.g., \**huwwa ma-bi-ya-ḏhab-š* instead of *huwwa ma-bi-yi-ruḥ-š* “he does not go”).

Production tokens of the present negation were coded as completely incorrect if they involved the following: correct usage of verbal present negators with verbs inflected for both the wrong tense and agreement (e.g., \**hiya lā ḏahaba* instead of *hiya lā ta-ḏhab* “she does not go”); unauthentic usage of negators with verbs inflected for either the wrong agreement or tense (e.g., \**hiya lā yi-rūḥ* instead of *hiya ma-bi-t-ruḥ-š* “she does not go”); correct and authentic usage of negators with lack of Arabic verb suppliance (e.g., using an English verb instead of its Arabic equivalent); correct and authentic usage of negators with verbs inflected for the wrong agreement (e.g., \**hiya lā ya-ḏhab* instead of *hiya lā ta-ḏhab* “she does not go”); correct and authentic usage of negators with verbs inflected for the wrong tense (e.g., \**huwa lā ḏahab* instead of *huwa lā ya-ḏhab* “he does not go”); correct usage of the EA negator *ma-š* “does/do not” with MSA-like verbs unpreceded by the habitual marker *bi-* (e.g., \**huwa ma-ya-ḏhab-š* instead of *huwwa ma-bi-yirūḥ-š* “he does not go”).

Recall, MSA and EA exhibit the negator *mā* and the non-discontinuous negator *muš*, respectively, to negate verbal present sentences. The latter was somewhat present in the data produced by the control participants. However, these two negators were not presented in the L2 participants’ instructional textbooks as negators of the verbal present tense. Accordingly, any

production tokens of the MSA negator *mā* with correctly inflected verbs were coded as completely incorrect due to the instructional input error (e.g., *huwa mā ya-ḏhab* “he does not go”). Similarly, any production tokens of the EA non-discontinuous negator *muš* with correctly inflected EA-like verbs were coded as completely incorrect due to the instructional input error (e.g., *huwwa muš bi-yi-rūḥ* “he does not go”).

### 3.6.2.5 Coding criteria for the subjunctive subordination data

Each data sample was expected to contain 12 obligatory-in-context tokens of the subjunctive subordination. Production tokens of the optional structure (i.e., the verbal noun instead of the *ʔan* construction) were not counted toward the overall production of the subjunctive subordination. However, those optional production tokens were coded separately to examine the extent to which the participants opted for the simple structure in comparison to its complex counterpart.

To determine the authenticity of the production tokens of the subjunctive subordination, whether they were in MSA or EA, the matrix and subordinate verbs were considered first. If either verb is exclusively used in MSA (e.g., *yušāhid* “to watch,” *yaḏhab* “to go”), then the MSA complementizer *ʔan* had to be overtly used. Similarly, if either of the produced verbs is exclusively used in EA (e.g., *yišūf* “to see/to watch” and *yirūḥ* “to go”), then the complementizer had to be obligatorily nulled, and the habituality prefix marker *bi-* had to be used with the matrix verb in the contexts of simple present tense and present progressive tense (i.e., its obligatorily used context). In contrast, if the matrix and subordinate verbs are mutually used in both varieties (e.g., *yanām* “to sleep,” *yasbaḥ* “to swim,” *yaʔkul* “to eat,” *yašrab* “to drink”), then the former was considered only. In the simple present and present progressive contexts, the habituality prefix *bi-* obligatorily precedes the matrix verb in EA. If the prefix is absent, then the whole production was treated as an MSA sentence. In this case, the MSA complementizer *ʔan* must be

overtly spelled out. If the matrix clause exhibited the EA-like active participles (e.g., *ḩāyiz/a* “wanting”), then the complementizer *ʔan* had to be obligatorily absent.

Failure in the application of the above-mentioned rules renders the whole construction unauthentic. Hence, if an EA-like production did not contain any morphosyntactic errors but involved lack of suppliance for the habitual marker *bi-*, then it was coded as partially incorrect due to the unauthentic nature of the production (e.g., *\*huwwa yi-fakkar yi-rūḩ* as opposed to *huwwa bi-yi-fakkar yi-rūḩ* “he is thinking to go”). If the complementizer *ʔan* “to” was spelled out in an EA-like production, then it was coded as partially incorrect (e.g., *\*hiyya ḩāyiz-a ʔan ta-ʔkul* as opposed to *hiyya ḩāyiz-a tā-kul* “she wants to eat”). Similarly, lack of suppliance of the complementizer *ʔan* “to” in a production containing mutually used verbs in both varieties was coded as partially incorrect because of the unauthentic nature of the production (e.g., *\*huwa yu-ḩibb ya-lḩab* as opposed to *huwa yu-ḩibb ʔan ya-lḩab* “he likes to play”).

### **3.6.2.6 Coding criteria for the cause–effect subordination data**

Each data sample was expected to contain a fixed number of eight obligatory-in-contexts tokens of the cause–effect subordination. Production tokens of the optional structure were not counted toward the overall production of the cause–effect subordination—that is, *bisabab* and *ḩašān/ḩalašān* “because of” if followed by an *ʔiḩāfah* phrase instead of a subordinate clause headed by the conjunctions *liʔanna* or *ḩašān/ḩalašān* “because.” Nonetheless, those optional production tokens were coded separately to examine the extent to which the participants opted for the simple structure in comparison to its complex counterpart.

Recall, both conjunctions *liʔanna* or *ḩašān/ḩalašān* “because” are used interchangeably in EA. Since both conjunctions were present in the L2 participants’ textbooks, it was necessary to develop certain criteria to determine the authenticity of the cause–effect subordination

production tokens, whether they were in MSA or EA. To elaborate, the matrix and subordinate verbs in any production tokens were considered first. If either one is exclusively used in MSA (e.g., *yušāhid* “to watch,” *yaḏhab* “to go”), then the MSA conjunction *liʔanna* “because” had to be used. If the EA conjunction *ʕašān/ʕalašān* “because” was used instead, then the whole production was considered unauthentic and hence partially incorrect due to the unauthentic sound of the production. Conversely, if either one of the verbs is exclusively used in EA, then usage of either one of the conjunctions was accepted. However, the verb in the matrix clause had to be preceded by the habitual prefix *bi-* in the contexts of simple present and present progressive. If not, then the whole production was coded as partially incorrect due to the unauthentic nature of the production. If the produced verbs are mutually used in both varieties (e.g., *yanām* “to sleep,” *yasbaḥ* “to swim,” *yaʔkul* “to eat,” *yašrab* “to drink”), then the verb in the matrix clause was considered only. If the EA habitual prefix *bi-* was absent in its obligatorily-used context, then the whole production tokens were treated as an MSA construction. Accordingly, the conjunction *liʔann* “because” was expected to be used. If the EA conjunction *ʕašān/ʕalašān* “because” was used instead, then the whole production was considered unauthentic and coded as partially incorrect.

Moreover, production tokens of the genitive construction *bisabab* “because of” followed by verbal sentences, nominal-verbless sentences, or nominal-verbal sentences were coded as completely incorrect (e.g., *\*hiya ḏahab-at ʔilā ʔad-duktūr bisabab kān-at marīḏa*, *\*hiya ḏahab-at ʔilā ʔad-duktūr bisabab hiya marīḏ-a*, *\*hiya ḏahab-at ʔilā ʔad-duktūr bisabab hiya kānat marīḏ-a* “\*she went to the doctor **because of she was sick**”). This is so because the genitive construction *bisabab* “because of” is always followed a genitive noun phrase only (e.g., *hiya ḏahab-at ʔilā ʔad-duktūr bisabab maraḏi-hā* “she went to the doctor because of her sickness”).

## 3.7 Data Analysis

### 3.7.1 Fluency

Fluency in L2 studies is typically gauged based on the principle of how fast L2ers access and produce linguistic units (e.g., syllabus, words, phrases, clauses, sentences) during a unit of time (e.g., second, minute, conversational session). Thus, the current study adopted the concept of speed rate as a predictor of the participants' fluency, which was measured by counting Arabic words produced per minute (W/M) in each of the two narrative tasks (as in Albirini, 2018, 2019; Raish, 2018). Certain lexemes were not counted, including English produced words, filler words and phrases (e.g., *yaʕnī* “meaning,” *ma ʔaʕraf-š*, *lā ʔaʕrif* “I do not know,” *lā ʔataðakkar* “I do not remember”), the conjunction *wa* “and,” the definite article *ʔal* “the,” and enquiries about the task pictures.

### 3.7.2 Accuracy

Accuracy was operationalized based on the participants' ability to produce output that is free of errors and that is native-like in nature. However, instead of gauging accuracy quantitatively at a macro-based level (as in Albirini, 2018, 2019; Raish, 2018), it was measured qualitatively and quantitatively at a micro-based level. The qualitative method was implemented to provide a fine-grained analysis of the correct and incorrect rule application in the production of each of the simple target structures of the past tense, present tense, past negation, and verbal present negation (i.e., simple T-units). This analysis can lay the basis for examining the role of input, L1 transfer, and implicit versus explicit learning in the acquisition of these structures by the L2 participants. Quantitatively, accuracy was measured by computing the ratio of the correct production tokens in each of these target structures to the total production tokens of each one of



them. The calculation process of each participant's accuracy on each of these target structures is illustrated in the following equation:

$$\frac{\text{Number of Correct Tokens of Each Simple Target Structure}}{\text{Number of Total Tokens of Each Simple Target Structure}} \times 100$$

### **3.7.3 Complexity**

The notion of subordination was adopted in the current study to measure the complexity construct in the participants' output, particularly in the production of subjunctive subordination and cause–effect subordination (i.e., complex T-units). Given that the data were elicited in obligatory contexts, the participants' complexity was evaluated based on their ability to produce error-free tokens only in these two complex structures. Precisely, the complexity of the participants was measured qualitatively and quantitatively. The qualitative method was implemented to provide a fine-grained analysis of the correct and incorrect rule application in the production of each of these two complex subordination structures. This analysis can again lay the ground to examine the role of input, L1 transfer, and implicit versus explicit learning in the acquisition of these structures by the L2 participants. Quantitatively, complexity was measured by means of calculating the ratio of the error-free production of the subjunctive subordination and cause–effect subordination structures to the total production tokens for each of these two structures. The calculation process of each participant's production of each of the above-mentioned two complex target structures is illustrated in the following equation:

$$\frac{\text{Number of Correct Tokens of Each Complex Structure}}{\text{Number of Total Tokens of Each Complex Structure}} \times 100$$

### **3.7.4 Overall Proficiency**

Following Albirini (2018, 2019), the overall proficiency of the participants was gauged by adding up the scores obtained in the three constructs of fluency, accuracy, and complexity.

### **3.8 Summary**

This chapter provided detailed discussions of the instruments employed in the present study. Specifically, it discussed the recruitment criteria and results of the participants, target structures, timing and frequency of the target structures in the L2 participants' textbooks, study tasks, data collection, coding, and analysis. Given that the primary aim of the study was to compare the learning outcomes of two L2 groups exposed to different methods of instruction (IPG and UIPG), it was indispensable to have a high level of homogeneity between the L2 groups at each proficiency level (first year, second year, and third year). Thus, various variables among the L2 participants had to be tightly controlled for, including their L1, prior exposure to Arabic, and proficiency-placement level. Accordingly, all 57 L2 participants representing the two L2 groups were English L1 speakers, were not heritage speakers of Arabic, had not resided in any Arabic-speaking countries, and had not studied in a previous Arabic L2 program for more than 6 months. Moreover, both the IPG and UIPG were compatible in terms of their respective proficiency-placement levels. To compare their performance with that of native speakers of Arabic, a CG consisting of nine native speakers of EA was recruited. The controls were doctoral and postdoctoral students, were L2 speakers of English, and had not lived in the U.S. for more than 5 years. The participants performed three picture-based tasks: two narrative tasks and one manipulation task. They were designed to elicit spontaneous obligatory-in-context data of the target structures of past tense, present tense, past negation, verbal present negation, subjunctive subordination, and cause-effect subordination. Importantly, the tasks contained actions that required using basic words to which the L2 participants were exposed at an early stage of their learning process. The data were elicited from the participants in the form of one-on-one interviews. The collected data were quantitatively analyzed based on the most widely adopted measures of the fluency, accuracy, and complexity constructs. Fluency was gauged by counting

Arabic W/M in each of the two narrative-based tasks. Accuracy was measured at a micro-based level by providing a qualitative analysis of the correct and incorrect use of the simple target structures of past tense, present tense, past negation, and present negation (i.e., simple T-units). It was also measured quantitatively by means of computing the ratio of the correct tokens of each of these simple target structures to the total production tokens of each of these structures. Complexity was likewise measured qualitatively by providing descriptive analyses of the correct and incorrect use of the complex target structures of subjunctive and cause–effect subordinations (i.e., complex T-units). It was also measured quantitatively by means of calculating the ratio of the correct tokens of each of these complex target structures to the total production tokens of each of these structures. The overall proficiency was measured by adding up the scores obtained in the three proficiency constructs of fluency, accuracy, and complexity. Detailed analyses and discussions of the data results are provided in Chapters 4 and 5 below.

## Chapter 4 Results

### 4.1 Overview

As discussed in Chapter 3, the current study aimed to compare the learning outcomes of Arabic L2ers representing two L2 groups: the integrated program group (IPG) and unintegrated program group (UIPG) across three proficiency levels: first year, second year, and third year. In addition, the performance of the two L2 groups was compared to that of a control group (CG) consisting of nine native speakers of Egyptian Arabic (EA). Spontaneous production data were elicited from all participants, who performed three picture-based tasks (two narrative tasks and one manipulation task) through in-person one-on-one interviews. This chapter provides an analysis of the findings of the elicited data by answering the first research question pertinent to the nature of language output (in terms of fluency, accuracy, complexity, and overall proficiency) of the IPG versus that of UIPG.

Overall, 11 dependent variables were examined: (a) fluency in the past tense context, (b) fluency in the present tense context, (c) accuracy on the past tense with the 3.s.m agreement form, (d) accuracy on the past tense with the 3.s.f agreement form, (e) accuracy on the present tense with the 3.s.m agreement form, (f) accuracy on the present tense with the 3.s.f agreement form, (g) accuracy on the past negation, (h) accuracy on verbal present negation, (i) complexity in the production of subjunctive subordination, (j) complexity in the production of cause–effect subordination, and (k) overall proficiency.

The production data were analyzed qualitatively and quantitatively. The former was relied on to provide a full scope of how the target morphosyntactic structures developed in the L2 participants' interlanguage systems. Thus, qualitatively delineating analyses, accompanied by an ample number of samples extracted from the data, were provided to examine how the grammatical rules of the target structures were correctly and incorrectly applied by the IPG and UIPG across all proficiency levels.

The quantitative analysis of the data entailed providing descriptive statistics for the overall raw data, as well as boxplot and line figures for the mean scores of all participants representing each proficiency level. Moreover, the above-mentioned 11 dependent variables were anticipated to be influenced by two fixed factors. The first was the L2 participant group factor (two groups: IPG and UIPG), which was determined as an independent between-group variable. The second was the L2 proficiency level factor (three levels: first year, second year, and third year), which was determined as an independent within-group variable. Therefore, multiple full factorial two-way analyses of variance (ANOVA) tests were conducted in SPSS to examine the main and combined effects of the two independent variables of the L2 participant group and L2 proficiency level on each of the 11 continuous dependent variables mentioned above.

The control participants were not included in the two-way ANOVA analyses for two reasons. First, unlike the two L2 groups (IPG and UIPG) that corresponded to three proficiency levels (first year, second year, and third year), the CG corresponded to only one proficiency level (native). As a consequence, including the CG in this statistical model would dilute its purpose and would likely yield spurious findings. Second, because the control participants outperformed the L2 participants in almost all measures, with very large variances in some of the dependent variables examined, adding the former to the two-way ANOVA analyses would certainly

introduce additional noise to this model. It could consequently result in violating the assumption of homogeneity of variance, which is considered a vital assumption for parametric tests to function properly. Violation of this assumption would thus increase the occurring chances of Type I error—that is, false positive or false conclusion that the results are significant when they are actually not (Tabachnick & Fidell, 2007). Alternatively, secondary analyses were performed to compare the performance of each proficiency level within each L2 group with that of the CG. When conducting those alternative comparison analyses, two independent variables were determined: the control participants and the L2 participants representing each level within each L2 group. Therefore, independent samples *t*-tests were performed.

In all statistical tests, an alpha of .05 was utilized as an indicator for the significance level. For brevity purposes, the main results revealed by the two-way ANOVA tests were fully written in the narrative prose, but all follow-up comparisons were included as tables in the Appendices section. As for the results revealed by the independent samples *t*-tests, they were summarized in tables incorporated into the narrative prose.

In the subsequent sections, the results obtained from the participants' production data are discussed in the following order: fluency, accuracy, complexity, and overall proficiency.

## **4.2 Fluency**

Fluency, as the first proficiency construct, was examined in the two narrative tasks for the past tense and present tense. Examination of fluency was based on the concept of speed rate as a predictor of the fluency of the participants' oral production. It was gauged by calculating Arabic words produced per minute (W/M) in each of the two narrative tasks. Completion of the narrative tasks was self-paced, as the participants completed the tasks at their own pace. Table 4.1 provides descriptive statistics for the fluency means of the IPG, UIPG, and CG in the past

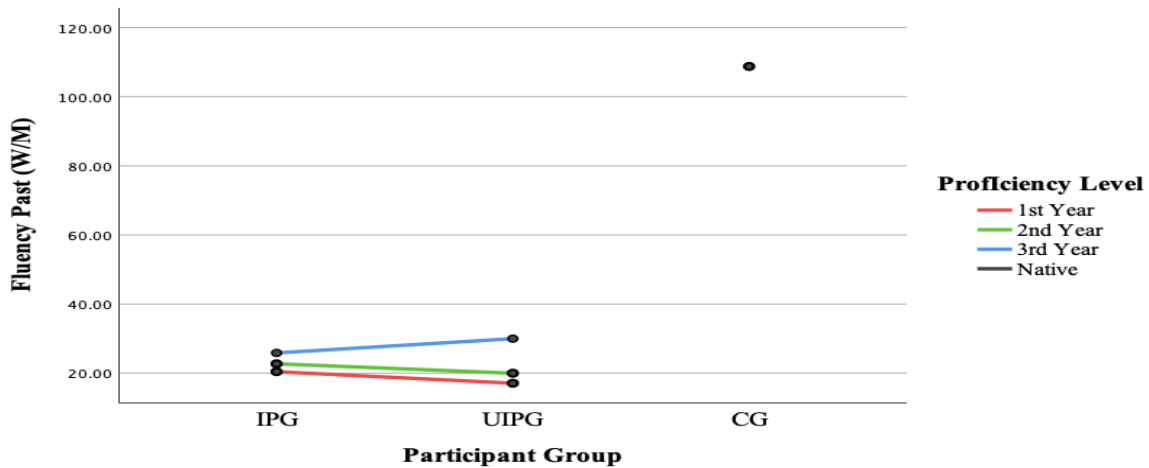
and present tense narrative tasks (see also Figures 4.1 and 4.2 that visualize the fluency trends among all participants in both tasks). It is evident that when narrating past and present events, the fluency of both L2 groups (IPG and UIPG) increased somewhat as their proficiency increased—from first year to second year to third year. Additionally, the W/M means of the IPG participants across all proficiency levels were relatively comparable to those of their UIPG counterparts. Furthermore, the IPG and UIPG across all proficiency levels, as well as the CG, produced W/M in the present tense task at a slightly higher speed than they did in the past narrative task. Expectedly, the W/M means of the CG were substantially higher than both L2 groups across all proficiency levels in both narrative tasks.

**Table 4.1** *Descriptive Statistics for the Fluency of the IPG, UIPG, and CG in the Past and Present Tense Narrative Tasks*

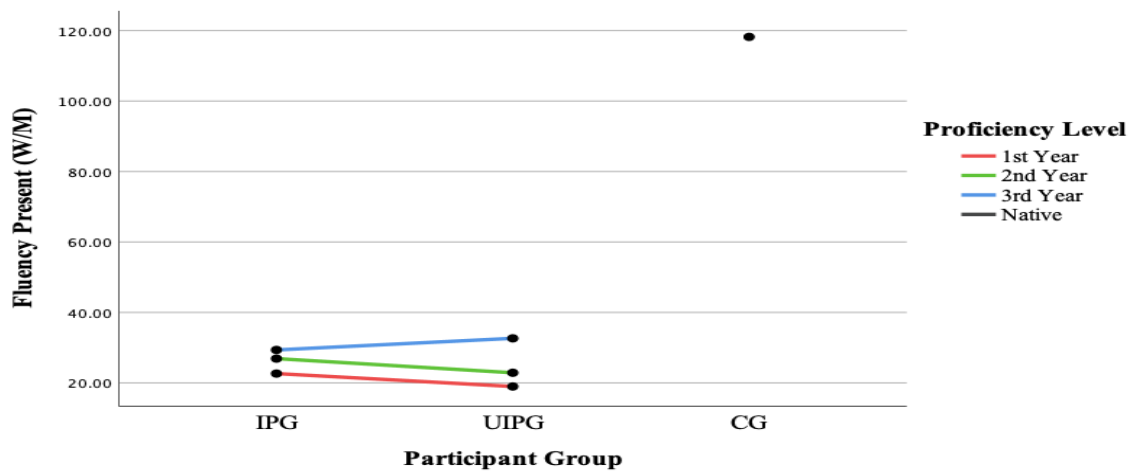
Participant Group Proficiency Level	Past Task Fluency			Present Task Fluency		
	Total W/M	<i>M</i>	<i>SD</i>	Total W/M	<i>M</i>	<i>SD</i>
<b>IPG</b>						
1st year ( <i>n</i> = 9)	184	20.44	7.70	204	22.67	8.86
2nd year ( <i>n</i> = 11)	250	22.73	9.59	296	26.91	9.83
3rd year ( <i>n</i> = 8)	207	25.88	8.53	235	29.38	8.96
<b>UIPG</b>						
1st year ( <i>n</i> = 10)	171	17.10	3.35	190	19.00	4.22
2nd year ( <i>n</i> = 10)	200	20.00	7.21	229	22.90	6.23
3rd year ( <i>n</i> = 9)	270	30.00	9.00	294	32.67	8.67
<b>CG</b>						
Native ( <i>n</i> = 9)	979	108.79	7.81	1064	118.22	7.81

*Note.* W/M = Words produced per Minute; M = Mean; SD = Standard Deviation.

**Figure 4.1** Fluency (W/M) Means of the IPG, UIPG, and CG in the Past Tense Narrative Task



**Figure 4.2** Fluency (W/M) Means of the IPG, UIPG, and CG in the Present Tense Narrative Task



Two full factorial two-way ANOVA tests were conducted to examine the main effects of the L2 participant group and the L2 proficiency level factors, as well as their interaction effect, on the fluency W/M means of the two L2 groups across the three proficiency levels in the past



and present narrative tasks. The findings of the first two-way ANOVA test performed on the L2 participants' fluency W/M means in the past tense task revealed several findings. The main effect of the L2 participant group factor was statistically not significant ( $F(1,51) = .097, p = .757$ , partial  $\eta^2 = .002$ ). The main effect of the L2 proficiency level factor was statistically significant ( $F(2,51) = 6.466, p = .003$ ) and yielded an effect size of .202, indicating that 20.2% of the variances in the L2 participants' fluency were explained by their proficiency levels. However, the interaction effect between the two variables was not significant ( $F(2,51) = 1.248, p = .296$ , partial  $\eta^2 = .047$ ), indicating that there was no combined effect of the L2 participant group and L2 proficiency level factors on the L2 participants' fluency when narrating past tense events. To reveal the nature of differences between all L2 proficiency levels, Tukey's post hoc analyses were conducted (Table 1f: Appendix F). The findings revealed that the first-year and second-year participants did not differ significantly from one another ( $p = .514$ ). However, the third-year participants differed significantly from both the first-year participants ( $p = .002$ ) and second-year participants ( $p = .032$ ). Precisely, based on the follow-up pairwise comparisons for the W/M means of the three L2 proficiency levels within each of the L2 groups, the main effect of the L2 proficiency level factor was significant because the third-year UIPG participants produced W/M at a significantly higher speed than the first-year UIPG participants ( $p < .001$ ) and second-year UIPG participants ( $p = .008$ ), whereas the W/M means of the IPG were statistically comparable across all proficiency levels (Table 2f: Appendix F).

As for the second two-way ANOVA test performed on L2 participants' fluency W/M means in the present tense narrative task, it revealed identical results to those stated above. That is, the main effect of the L2 participant group factor was statistically not significant ( $F(1,51) = .470, p = .496$ , partial  $\eta^2 = .009$ ). The main effect of the L2 proficiency level factor was

statistically significant ( $F(2,51) = 7.300, p = .002$ ) and yielded an effect size of .223, indicating that 22.3% of the variances in the L2 participants' fluency when narrating events in the present tense were explained by their proficiency levels. However, the interaction effect between the two variables was not significant ( $F(2,51) = 1.184, p = .314, \text{partial } \eta^2 = .044$ ), indicating that there was no combined effect of the L2 participant group and L2 proficiency level factors on the L2 participants' fluency when narrating present tense events. Tukey's post hoc tests (Table 3f: Appendix F) revealed that the first-year and second-year participants did not differ significantly from one another ( $p = .222$ ). However, the third-year participants differed significantly from the first-year participants ( $p < .001$ ) and near-significantly from the second-year participants ( $p = .059$ ). Precisely, the follow-up pairwise comparisons for the W/M means of the three L2 proficiency levels within each of the L2 groups revealed that the main effect of the L2 proficiency level factor was significant because the third-year UIPG participants produced W/M at a significantly higher speed than the first-year UIPG participants ( $p < .001$ ) and second-year UIPG participants ( $p = .011$ ), whereas the W/M means of the IPG did not improve significantly as their proficiency increased from the first-year to second-year to third-year level (Table 4f: Appendix F).

To examine whether the control participants differed from the L2 participants representing each proficiency level within each L2 group, 12 independent sample *t*-tests were performed: six on the fluency W/M means in the past tense task and six on the fluency W/M means in the present tense task. The statistical analyses showed that the CG produced W/M at a significantly faster rate than the L2 IPG and UIPG participants across all proficiency when narrating both past tense events (Table 4.2) and present tense events (Table 4.3).

**Table 4.2** Results of *t*-tests on the Fluency W/M Means of the CG and Each Proficiency Level Within Each L2 Group in the Past Tense Narrative Task

Variable (I)	Variable (J)	<i>df</i>	<i>t</i>	Sig. <i>p</i>	95% Confidence Interval	
					Lower Bound	Upper Bound
CG	IPG					
Native ( <i>n</i> = 9)	1st year ( <i>n</i> = 9)*	9.708	-10.745	< .001	-106.73	-69.94
	2nd year ( <i>n</i> = 11)*	10.192	-10.332	< .001	-104.56	-67.54
	3rd year ( <i>n</i> = 8)*	10.300	-9.903	< .001	-101.48	-64.32
CG	UIPG					
Native ( <i>n</i> = 9)	1st year ( <i>n</i> = 10)*	8.294	-11.632	< .001	-109.74	-73.62
	2nd year ( <i>n</i> = 10)*	9.362	-10.912	< .001	-107.07	-70.48
	3rd year ( <i>n</i> = 9)*	10.311	-9.416	< .001	-97.34	-60.21

Note. \* = Equality of variances was not assumed via Levene's *F* test (*p* < .05).

**Table 4.3** Results of *t*-tests on the Fluency W/M Means of the CG and Each Proficiency Level Within Each L2 Group in the Present Tense Narrative Task

Variable (I)	Variable (J)	<i>df</i>	<i>t</i>	Sig. <i>p</i>	95% Confidence Interval	
					Lower Bound	Upper Bound
CG	IPG					
Native ( <i>n</i> = 9)	1st year ( <i>n</i> = 9)**	16	-16.416	< .001	-107.90	-83.22
	2nd year ( <i>n</i> = 11)**	18	-16.352	<.001	-103.05	-79.58
	3rd year ( <i>n</i> = 8)**	15	-14.536	<.001	-101.88	-75.82
CG	UIPG					
Native ( <i>n</i> = 9)	1st year ( <i>n</i> = 10)*	9.130	-19.117	<.001	-110.94	-87.51
	2nd year ( <i>n</i> = 10)*	10.435	-17.690	<.001	-107.26	-83.38
	3rd year ( <i>n</i> = 9)**	16	-14.777	<.001	-97.83	-73.28

Note. \* = Equality of variances was not assumed via Levene's *F* test (*p* < .05); \*\* = Equality of variances was assumed via Levene's *F* test (*p* > .05).

To sum up the findings revealed by the statistical analyses altogether, the L2 participant group factor (IPG and UIPG) did not significantly affect the L2ers' fluency when narrating events in either the past tense or present tense because the IPG produced W/M at a significantly comparable rate to W/M produced by the UIPG at each proficiency level. Only the L2 proficiency level factor yielded significant main effects on the L2 participants' fluency.

However, as revealed by the follow-up within-group pairwise analyses, the significance was mainly due to outperformance of the third-year UIPG participants, whose fluency's W/M mean scores were significantly higher than those of the first-year and second-year UIPG participants, contrary to the IPG, whose W/M means were statistically comparable across all proficiency levels in both narrative tasks for the past and present tenses. This, in turn, can be argued to provide an advantage for the UIPG over the IPG because the fluency of the former improved significantly over time, particularly at the third-year level, compared to the latter, whose fluency did not show any statistically significant improvement in the long run. Notwithstanding, the fluency of the CG was significantly higher than that of both L2 groups across all proficiency levels in both narrative tasks.

### **4.3 Accuracy**

This section presents the results related to the accuracy construct of all participants in the IPG, UIPG, and CG. This proficiency dimension was examined particularly in the production data of the structures of past tense, present tense, past negation, and verbal present negation in the 3.s.m and 3.s.f contexts, which were elicited in the two narrative tasks (see Section 3.4.2.1: Chapter 3). The results revealed by the tense and negation data are discussed below in sections 4.3.1 and 4.3.2, respectively.

#### **4.3.1 Tense**

##### ***4.3.1.1 Past tense***

Based on the coding criteria discussed in Section 3.6.2.1, the production data of the past tense were analyzed accordingly. Table 4.4 provides descriptive statistics for the production accuracy of the IPG, UIPG, and CG on the past tense inflected for the 3.s.m and 3.s.f agreement forms.



- (40) a. *hiya jar-at mumkin fī ʔaṣ-ṣabāḥ*  
 she Perf.ran-3.s.f probably in the-morning  
 “She ran probably in the morning.”  
 (IPG: 1st-year participant)
- b. *hiya sāfar-at bi ʔas-safīn-a*  
 she Perf.traveled-3.s.f by the-ship-s.f  
 “She traveled by ship.”  
 (IPG: 2nd-year participant)
- c. *hiya ʔištar-at malābis*  
 she Perf.bought-3.s.f cloths  
 “She bought cloths.”  
 (UIPG: 2nd-year participant)
- d. *rajaʕ-at ʔal-marʔa ʔilā bayt-i-hā*  
 Perf.returned-3.s.f the-woman to house-Gen-her  
 “The woman returned to her house.”  
 (UIPG: 3rd-year participant)

The incorrect tokens produced by the participants in the 3.s.m past tense context were categorized into four patterns of errors, as summarized in Table 4.5. The most frequent error pattern was labeled as lack of suppliance, which constituted 36.9% of the total errors produced. It occurred when the participants produced English verbs instead of their Arabic equivalents, as in (41a–b); produced verbal noun forms instead of their past counterparts, as in (41c–d); and avoided the whole structure altogether. The second most frequent error was due to inflecting the verbs for the wrong tense, and it constituted 34.7% of the total errors. It occurred mainly when present tense forms were used instead of their past counterparts, as in (42a–b). The third most frequent error was due to inflecting the verbs for both the wrong tense and agreement, and it constituted 16.4% of the total errors. It occurred when the L2 participants produced present forms instead of their past counterparts and inflected them for the 1.s or 3.s.f forms instead of the 3.s.m form, as in (43a–b), respectively. The fourth and least frequent pattern of errors was due to







b. IL: <i>*huwa</i>	<i>šāhad-t</i>	<i>ʔas-sīnimā</i>
he	Perf.watched-1.s	the-theater
TL: <i>huwa</i>	<i>šāhada</i>	<i>ʔal-ʔaflām</i>
he	Perf.watched.3.s.m	the-movies
“He watched the movies.”		

(UIPG: 2nd-year participant)

(45) CG: <i>*bi-yi-sbaḥ</i>	<i>fī</i>	<i>mayy-a</i>
Hb-Imperf.3.s.m-swim	in	water-f
TL: <i>sabaḥ</i>	<i>fī</i>	<i>mayy-a</i>
Perf.swam.3.s.m	in	water-f
“He swam in the water tea.”		

(control participant)

As for the incorrect tokens produced by the participants in the 3.s.f past tense context, the qualitative analysis revealed six patterns of errors, as summarized in Table 4.6. The most prevalent error pattern was due to inflecting the produced past verbs for the wrong agreement, and it constituted 36.7% of the total errors. This error occurred when the verbs were inflected for the 3.s.m, 1.s, 2.s.f, and 2.s.m forms instead of the 3.s.f form, as in (46a, b, c, d), respectively. The second most frequent error was due to the production of present tense forms instead of their past tense counterparts, as in (47a–b), and it constituted 29% of the total errors. The third most frequent pattern was labeled as lack of suppliance, constituting 24.6% of the total errors. It occurred when the participants produced English verbs instead of their Arabic equivalents, as in (48a); produced verbal noun forms instead of past forms, as in (48b); produced nominal-verbless present sentences, as in (48c); and avoided describing some pictures altogether. The fourth most frequent error occurred when the verbs were inflected for both the wrong tense and agreement, as in (49a–b), and it constituted 8.3% of the total errors. The fifth type of errors constituted only 0.7% of the total incorrect tokens and was produced by two participants in the IPG and UIPG, when they used the referent subject pronoun *huwa* “he” instead of *hiya* “she,” as in (50). The last



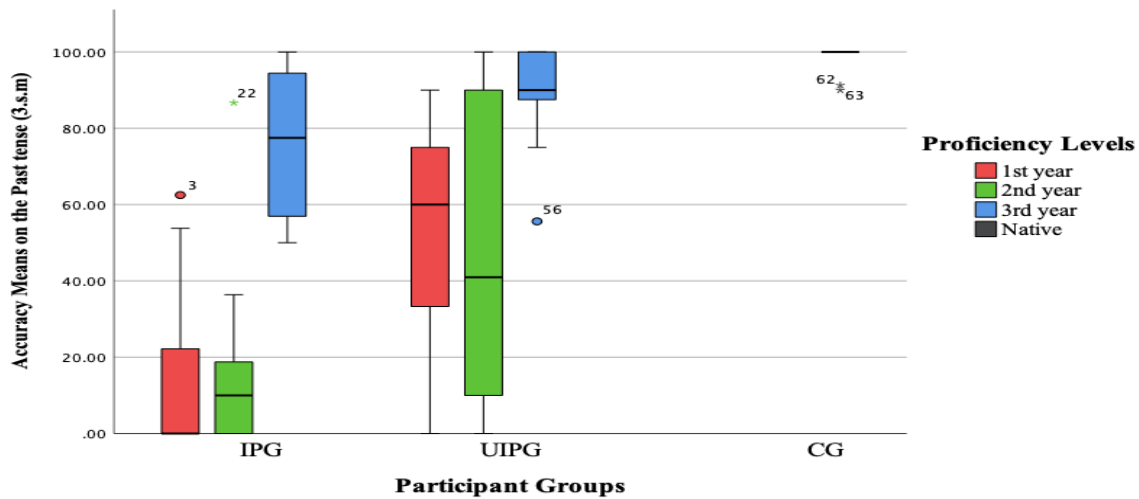




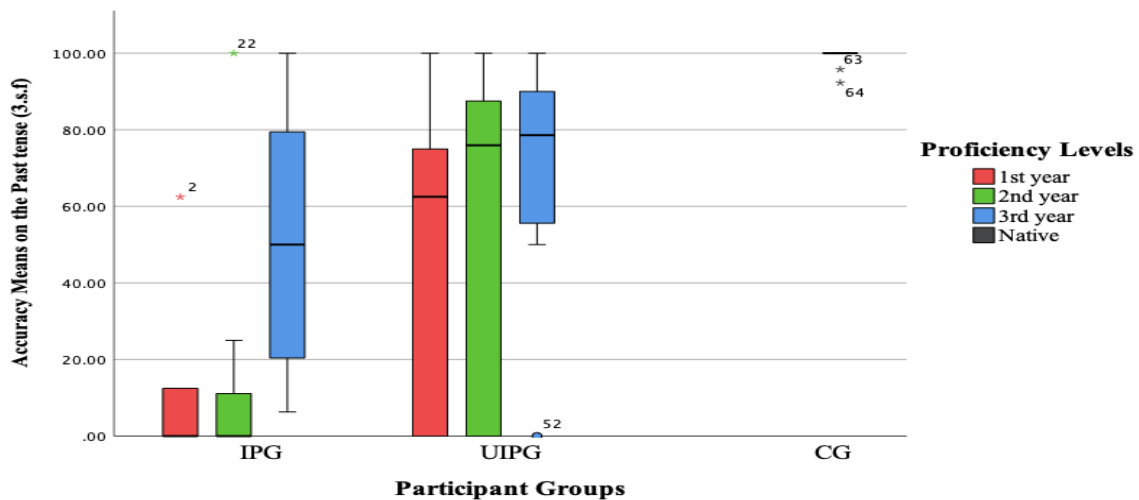
b.	<i>kān</i>	<i>bi-yi-rtāḥ</i>	<i>min taslluq</i>	<i>ʔil-gibāl</i>
	Perf.was.3.s.m	Prog-Imperf.3.s.m-rest	from climbing.VN	the-mountain
	“He was resting from climbing the mountain.”			
	(control participant)			
(52)	CG: <i>*mixaṭṭa-a</i>		<i>ti-lṣab</i>	tennis
	Planning-s.f		Imperf.3.s.f-play	tennis
	TL: <i>kān-it</i>	<i>mixṭṭa-a</i>	<i>ti-lṣab</i>	tennis
	Perf.was.3.s.f	planning-s.f	Imperf.3.s.f-play	tennis
	“She was planning [to] play tennis.”			
	(control participant)			

The aggregate accuracy data show certain degrees of variability in the two L2 groups’ production of the past tense in both the 3.s.m and 3.s.f contexts. As shown in Table 4.4 above, overall and except for the second-year UIPG participants, the IPG and UIPG participants across all proficiency levels exhibited a relatively higher accuracy rate in their use of the past tense in the 3.s.m context, compared to their use of this structure in the 3.s.f context. This indicates that the L2 participants had some difficulties in processing and producing the past tense with the 3.s.f agreement form, and this is supported by the qualitative analysis of the error patterns produced. That is, while the agreement error occurred the least in the 3.s.m context, it was the most frequent error produced in the 3.s.f context. Additionally, the accuracy rate of the UIPG participants was higher than that of their IPG counterparts across the three proficiency levels. This is evident visually in Figures 4.3 and 4.4, which show the trends (in quartiles) of the production accuracy of the IPG, UIPG, and CG on the past tense inflected for the 3.s.m and 3.s.f agreement forms, respectively.

**Figure 4.3** Boxplot for the Accuracy of the IPG, UIPG, and CG on the Past Tense with the 3.s.m Agreement Form



**Figure 4.4** Boxplot for the Accuracy of the IPG, UIPG, and CG on the Past Tense with the 3.s.f Agreement Form



As depicted in both figures, the interquartile ranges (i.e., the shaded second and third quartiles—the medial 50% of the scores) of the UIPG were higher than those of the IPG across all proficiency levels. Additionally, in the 3.s.m context, the accuracy rate increased slightly

from the first-year to second-year levels (especially within the UIPG), but it considerably increased at the third-year level within both L2 groups. However, in the 3.s.f context, the accuracy rate marginally increased as proficiency increased within the UIPG, and it did not increase at all from the first-year to second-year levels within the IPG, as evidenced by their accuracy medians (i.e., the line separating the shaded interquartile range) that were remarkably at the zero level. However, the accuracy rate substantially improved at the third-year level within the IPG. These observations point out that the L2 participant group, as a between-group factor, was in favor of the UIPG over the IPG and that the proficiency level, as a within-group factor, played an important role in the production accuracy on the past tense.

Moreover, and equally important, it is visually evident that the accuracy of the first-year and second-year IPG participants on the past tense with both agreement forms was clustered at the bottom of the two boxplots, compared to their UIPG counterparts, who exhibited greater variability in their accuracy. The within-group variability in the IPG is mainly observed among its third-year participants only, whose accuracy on both agreement forms became dispersed. Thus, the bottom clustered interquartile accuracy of the first-year and second-year IPG participants may indicate that learning did not take place among them, compared to their UIPG counterparts, whose widespread production accuracy signifies that learning was in progress. As for the third-year UIPG participants, their accuracy on the past tense with the 3.s.m form was tightly clustered at the top, strongly indicating that they reached a high level of accurate automatization. Still, their dispersed accuracy on the past tense in the 3.s.f context suggests that their learning was still in progress and that they had not fully automatized the past tense with the 3.s.f agreement form. One last important observation is that the accuracy of many participants of both L2 groups was at the ceiling level (100% of accuracy) in both agreement contexts. This

observation strongly suggests that ultimate attainment is possible in the acquisition of the past tense with the 3.s.m and 3.s.f agreement forms.

The above-mentioned observations of the two boxplots were confirmed by the two full factorial two-way ANOVA tests that were conducted to examine the main effects of the two independent variables (the L2 participant group and L2 proficiency level), as well as their interaction effects, on the accuracy of the two L2 groups across the three proficiency levels on the past tense in the 3.s.m and 3.s.f contexts. The first two-way ANOVA test that was performed on the production accuracy on the past tense inflected for 3.s.m revealed a number of findings. The main effect of the L2 participant group factor was significant ( $F(1,51) = 13.262, p < .001$ ) and yielded an effect size of .206, indicating that 20.6% of the variances in the L2 participants' accuracy were explained by the participant groups to which they belonged. The main effect of the proficiency level factor was likewise significant ( $F(2,51) = 20.044, p < .001$ ) and yielded an effect size of .440, indicating that 44% of the variances in the L2 participants' accuracy were explained by their proficiency levels. However, the interaction effect was not significant ( $F(2,51) = 1.006, p = .373, \text{partial } \eta^2 = .038$ ), indicating that there was no combined effect of the L2 participant group and L2 proficiency level factors on the L2 participants' accuracy on the past tense inflected for 3.s.m. To examine the nature of the differences between the three L2 proficiency levels, Tukey's post hoc analyses were conducted (Table 1g: Appendix G). The first-year and second-year participants did not differ significantly from one another ( $p = .854$ ). However, the third-year participants differed significantly from both the first-year participants ( $p < .001$ ) and second-year participants ( $p = .001$ ). Based on the follow-up pairwise comparisons between the IPG and UIPG at each proficiency level (Table 2g: Appendix G), the main effect of the L2 participant group factor was significant due to the significant outperformance of the first-



year UIPG participants over their IPG counterparts ( $p = .004$ ) and the second-year UIPG participants over their IPG counterparts ( $p = .021$ ). In addition, the follow-up pairwise comparisons between all proficiency levels within each L2 group revealed that the main effect of the L2 proficiency level factor was significant due to the outperformance of the third-year participants over the first-year and second-year participants within both the IPG and UIPG (Table 3g: Appendix G).

As for the second two-way ANOVA test that was performed on the production accuracy on the past tense inflected for 3.s.f, it revealed that the main effect of the L2 participant group factor was significant ( $F(1,51) = 11.369, p = .001$ ) and yielded an effect size of .186, indicating that 18.6% of the variances in the L2 participants' accuracy were explained by the L2 participant groups to which they belonged. The main effect of the L2 proficiency level factor was likewise significant ( $F(2,51) = 4.674, p = .014$ ) and yielded an effect size of .155, indicating that 15.5% of the variances in the L2 participants' accuracy were explained by their proficiency levels. However, the interaction effect was not significant ( $F(2,51) = .395, p = .675, \text{partial } \eta^2 = .015$ ), indicating that there was no combined effect of the L2 participant group and L2 proficiency level factors on the L2 participants' accuracy on the past tense inflected for the 3.s.f form. Tukey's post hoc analyses (Table 4g: Appendix G) revealed that the first-year and second-year participants did not differ significantly from one another ( $p = .977$ ). However, the third-year participants differed significantly from both the first-year participants ( $p = .020$ ) and second-year participants ( $p = .028$ ). Based on the follow-up pairwise comparisons between the IPG and UIPG at each proficiency level (Table 5g: Appendix G), the main effect of the L2 participant group factor was significant due to the outperformance of the UIPG over the IPG at the first-year level ( $p = .024$ ) and second-year level ( $p = .016$ ). In addition, the follow-up pairwise comparisons

between all proficiency levels within each L2 group yielded that the main effect of the proficiency level factor was significant due to the outperformance of the third-year participants over the first-year participants ( $p = .018$ ) and second-year participants ( $p = .023$ ) within the IPG only, whereas the accuracy means of the UIPG participants were statistically comparable across all proficiency levels (Table 6g: Appendix G).

To examine whether the control participants differed from the L2 participants representing each proficiency level within each L2 group, 12 independent sample  $t$ -tests were performed: six were conducted on the accuracy means of the past tense in the 3.s.m context and the other six were conducted on the accuracy means of the past tense in the 3.s.f context. The results revealed that, among all L2 participants, only the third-year UIPG participants were statistically comparable with the CG in their accuracy on the past tense in the 3.s.m context only (Table 4.7). However, the CG significantly outperformed both L2 groups across all proficiency levels in the 3.s.f context (Table 4.8).

**Table 4.7** Results of  $t$ -Tests on the Accuracy Mean Scores of the CG and Each Proficiency Level Within Each L2 Group on the Past Tense in the 3.s.m Context

Variable (I)	Variable (J)	$df$	$t$	Sig. $p$	95% Confidence Interval	
					Lower Bound	Upper Bound
CG	IPG					
Native ( $n = 9$ )	1st year ( $n = 9$ )*	8.424	-9.621	< .001	-102.14	-62.92
	2nd year ( $n = 11$ )*	10.628	-10.228	< .001	-98.01	-63.18
	3rd year ( $n = 8$ )*	7.544	-3.078	.016	-38.57	-5.33
CG	UIPG					
Native ( $n = 9$ )	1st year ( $n = 10$ )*	9.376	-4.642	.001	-66.36	-23.05
	2nd year ( $n = 10$ )*	9.245	-4.413	.002	-79.23	-25.67
	3rd year ( $n = 9$ )*	9.211	-1.837	.099	-21.21	2.16

Note. \* = Equality of variances was not assumed via Levene's  $F$  test ( $p < .05$ ).

**Table 4.8** Results of *t*-Tests on the Accuracy Mean Scores of the CG and Each Proficiency Level Within Each L2 Group on the Past Tense in the 3.s.f Context

Variable (I)	Variable (J)	df	t	Sig. p	95% Confidence Interval	
					Lower Bound	Upper Bound
CG	IPG					
Native (n = 9)	1st year (n = 9)*	8.291	-12.890	< .001	-104.77	-73.14
	2nd year (n = 11)**	18	-8.504	< .001	-106.37	-64.28
	3rd year (n = 8)*	7.080	-3.925	.006	-76.73	-19.12
CG	UIPG					
Native (n = 9)	1st year (n = 10)*	9.099	-4.197	.002	-80.41	-24.15
	2nd year (n = 10)*	9.078	-3.392	.008	-79.50	-15.93
	3rd year (n = 9)*	8.121	-2.678	.028	-52.96	-4.02

Note. \* = equality of variances was not assumed via Levene's *F* test ( $p < .05$ ); \*\* = equality of variances was assumed via Levene's *F* test ( $p > .05$ ).

To sum up, the qualitative and quantitative analyses of the participants' acquisition data for the past tense structure revealed a number of findings. Initially, the error that was due to producing present tense forms instead of their past counterparts was prevalent in the data of the first-year and second-year participants, but it was rarely produced by the third-year participants of both L2 groups. Thus, the crucial indication of this finding is that, despite the presence of the past tense feature in the participants' L1 (English), the participants at the beginner and intermediate levels exhibited a tendency to converge somewhat in their use of the present tense as the default form when narrating past tense events. Another finding is that both the IPG and UIPG were more accurate in their production of the past tense in the 3.s.m context than in the 3.s.f context, signifying some difficulties encountered by the L2 participants in processing and producing this tense with the 3.s.f agreement form. This is strongly supported by the fact that, while the agreement error was the least frequent error type in the 3.s.m context, it was the predominantly produced error in the 3.s.f context and was prevalently persistent across all proficiency levels of both L2 groups.

Furthermore, the first-year and second-year UIPG participants significantly outperformed their IPG counterparts in their accuracy on the past tense in both the 3.s.m and 3.s.f contexts. However, the IPG and UIPG were statistically comparable in their accuracy on this structure with both agreement forms at the third-year level. Additionally, although the first-year and second-year participants within each L2 group performed comparably in both agreement contexts, the accuracy of the third-year IPG participants was significantly higher than that of the first-year and second-year IPG participants in both agreement contexts. As for the third-year UIPG participants, their accuracy was significantly higher than that of the first-year and second-year UIPG participants in the 3.s.m context but statistically comparable in the 3.s.f context. One may infer from this finding that the IPG was better than the UIPG because the former significantly improved over time (at the third-year level) in their accuracy on the past tense with both agreement forms, compared to the UIPG, whose accuracy statistically increased over time (at the third-year level) in the 3.s.m context only but not in the 3.s.f context. However, upon further examination of the boxplot in Figure 4.4 above, an opposite inference was reached. That is, the interquartile accuracy medians of the first-year and second-year UIPG participants on the past tense with the 3.s.f form were higher than that of the third-year IPG participants. The implication of this observation is that the development of this tense structure with the 3.s.f agreement form took place rapidly at an earlier stage among the UIPG participants, compared to their IPG counterparts.

Lastly, among all the L2 participants across all proficiency levels, only the third-year UIPG participants were statistically on par with the CG in their accuracy on the past tense in the 3.s.m context. However, both the IPG and UIPG across all proficiency significantly underperformed the CG in the 3.s.f context. This finding suggests that, when it comes to the

acquisition of the past tense structure, only the UIPG participants were able to gradually approach native-like proficiency over time (in the 3.s.m context only).

#### 4.3.1.2 Present tense

Based on the coding criteria discussed in Section 3.6.2.2, the production data of the present tense were analyzed accordingly. Table 4.9 provides descriptive statistics for the production accuracy of the IPG, UIPG, and CG on the present tense inflected for the 3.s.m and 3.s.f agreement forms.

**Table 4.9** Descriptive Statistics for the Accuracy of the IPG, UIPG, and CG on the Present Tense in the 3.s.m and 3.s.f Contexts

Participant Group Proficiency Level	Present 3.s.m				Present 3.s.f			
	C/T	%	<i>M</i>	<i>SD</i>	C/T	%	<i>M</i>	<i>SD</i>
<b>IPG</b>								
1st year ( <i>n</i> = 9)	34.5/63	54.8	54.76	19.55	44.5/77	57.8	56.72	19.32
2nd year ( <i>n</i> = 11)	35.5/84	42.8	40.95	28.88	40.5/95	42.6	43.06	32.91
3rd year ( <i>n</i> = 8)	42/66	63.6	62.75	15.71	41.5/72	56.9	57.60	28.27
<b>UIPG</b>								
1st year ( <i>n</i> = 10)	41/74	55.4	55.06	20.08	39/80	47.6	47.36	24.04
2nd year ( <i>n</i> = 10)	41/72	56.9	57.61	27.25	47/82	58.8	57.50	24.40
3rd year ( <i>n</i> = 9)	59/68	86.8	86.37	16.11	72/86	83.7	84.19	12.84
<b>CG</b>								
Native ( <i>n</i> = 9)	93/95	97.9	97.66	4.66	110/110	100	100	0.00

*Note.* C/T = Correct/Total; % of correct tokens; *M* = Mean; *SD* = Standard Deviation.

Examples of the L2 participants' correct production tokens for the present tense with the 3.s.m and 3.s.f forms are provided in (53a–d) and (54a–d), respectively. Recall that MSA and EA diverge in the present tense (see Section 3.2.1.2: Chapter 3). Therefore, the data of this structure were coded separately, based on the variety in which they were produced. Out of all present tense tokens produced by the IPG participants, only 12 tokens were correctly inflected for the agreement and tense features in the EA variety: nine tokens in the 3.s.m context and three tokens in the 3.s.f context. Remarkably, of all these tokens, only one token was completely correct, as

provided in sentence (53c), in which the verb was uttered with the habitual prefix *bi-*. As for the remaining 11 tokens, they were partially incorrect, as will be discussed below. The third-year participants produced none of these EA tokens, whereas 11 tokens were produced by second-year participants and one token by a first-year participant. As for the control participants, they produced the present tense with both agreement forms of 3.s.m and 3.s.f in the EA variety, as in (55a–b), respectively. In other words, they produced the present verbs with the habitual prefix *bi-* in its obligatorily used context.

- (53) a. *huwa ya-šrab laban wa ya-ʔkul*  
 he Imperf.3.s.m-drink buttermilk and Imperf.3.s.m-eat  
 "He drinks buttermilk and eats."  
 (IPG: 1st-year participant)
- b. *wa fī ʔal-masāʔ huwa ya-fʕal ʔat-tamrīn*  
 and in the-evening he Imperf.3.s.m-do the-exercise  
 "And in the evening, he works out."  
 (UIPG: 1st-year participant)
- c. *huwa bi-yi-štayil fī ʔal-ʕamal*  
 he Hb-Imperf.3.s.m-work in the-work  
 "He works."  
 (IPG: 2nd-year participant)
- d. *huwa ya-ʔkul ʔal-faṭūr*  
 he Imperf.3.s.m-eat the-breakfast  
 "He eats breakfast."  
 (UIPG: 2nd-year participant)
- (54) a. *tu-darris fī ʔal-gāmiʕ-a*  
 Imperf.3.s.f-teach in the-university-s.f  
 "She teaches at the university."  
 (IPG: 1st-year participant)

b. *wa baʕda ɔ̄ālika ta-ʔkul* pizza  
 and after that Imperf.3.s.f-eat pizza  
 “And then she eats pizza.”  
 (UIPG: 1st-year participant)

c. IL: *wa baʕda ɔ̄ālika tu-nawwim ʔilā ʔal-layl*  
 and after that Imperf.3.s.f-sleep to the-night  
 TL: *wa baʕda ɔ̄ālika ta-nām fī ʔal-layl*  
 and after that Imperf.3.s.f-sleep to the-night  
 “And then she sleeps at night.”  
 (IPG: 3rd-year participant)

d. *liʔnna-hā tu-faddil ʔal-qahw-a*  
 because-she Imperf.3.s.f-perfer the-coffee-f  
 “Because she prefers coffee.”  
 (UIPG: 3rd-year participant)

(55) a. *ʔis-sāʕ-a ʕašr-a bi-yi-rūḥ šuyḷ-u*  
 the-clock-f ten-f Hb-Imperf.3.s.m-go work-his  
 “At ten o’clock, he goes to his work.”  
 (control participant)

b. *ʔis-sāʕ-a sitt-a wi nuṣ bi-ti-ysil sinan-hā*  
 the-clock-f six-f and half Hb-Imperf.3.s.f-wash teeth-her  
 “At six-thirty, she brushes her teeth.”  
 (control participant)

The incorrect tokens produced by the participants in the 3.s.m present tense context were categorized into five patterns of errors, as summarized in Table 4.10. The most frequent pattern of errors was due to lack of suppliance, constituting 64.2% of the total errors. This type of errors occurred when the participants produced mostly English verbs instead of their Arabic equivalents, as in (56a–b); produced verbal noun forms instead of their present counterparts, as in (57a–b); and avoided the description of some pictures altogether. The second most frequent error was due to producing past tense forms instead of their present tense counterparts, as in (58a–b), and this error constituted 18.4% of the total errors. An interesting observation regarding

this error pattern is that it exhibited an increasing tendency as proficiency increased within the IPG, but it decreased considerably at the third-year level within the UIPG.

**Table 4.10** Error Patterns Produced by the Participants in the 3.s.m Present Tense Context (Error Type Tokens/Total Errors)

Error Type	IPG			UIPG			CG
	1st year (n = 9)	2nd year (n = 11)	3rd year (n = 8)	1st year (n = 10)	2nd year (n = 10)	3rd year (n = 9)	Native (n = 9)
L of S	21/29	29/51	13/24	27/33	19/31	6/9	0/2
T	4/29	6/51	10/24	2/33	7/31	2/9	2/2
Agr	1/29	8/51	1/24	4/33	3/31	1/9	0/2
Authenticity*	1/29	7/51	0/24	0/33	0/31	0/9	0/2
Agr + T	2/29	1/51	0/24	0/33	2/31	0/9	0/2

Note. \* = coded as partially incorrect; L of S = Lack of Suppliance; T = Tense; Agr = Agreement.

- (56) a. IL: \**huwa*            exercises  
           he                    exercises  
       TL: *huwa*            *ya-tamarran*  
           he                    Imperf.3.s.m-exercise  
           “He exercises/works out.”

(IPG: 3rd-year participant)

- b. IL: \**huwa*            gets up  
           he                    gets up  
       TL: *huwa*            *ya-ṣhū*  
           he                    Imperf.3.s.m-wake up  
           “He wakes up.”

(UIPG: 3rd-year participant)

- (57) a. IL: \**huwa*            *nawm*  
           he                    sleeping.VN  
       TL: *huwa*            *ya-nām*  
           he                    Imperf.3.s.m-sleep  
           “He sleeps.”

(IPG: 2nd-year participant)



- b. IL: \**huwa*            *julūs*  
           he                    sitting.VN  
 TL: *huwa*                *ya-jlis*  
           he                    Imperf.3.s.m-sit  
           “He sits.”

(UIPG: 2nd-year participant)

- (58) a. IL: \**huwa*            *ḍahaba*                *ʔilā*                *ʔal*    gym  
           he                    Perf.went.3.s.m        to                the    gym  
 TL: *huwa*                *ya-ḍhab*                *ʔilā*                *ʔan-nādī*  
           he                    Imperf.3.s.m-go        to                the-gym  
           “He goes to the gym.”

(IPG: 1st-year participant)

- b. IL: \**kānā*                                *yu-daxxin*  
           Perf.was.3.s.m                    Imperf.3.s.m-somke  
 TL: *yu-daxxin*  
           Imperf.3.s.m-somke  
           “He smokes.”

(UIPG: 3rd-year participant)

The third most frequent error was due to inflecting the produced present verbs for the wrong agreement, and it constituted 10.1% of the total errors. This error occurred when the verbs were inflected for the 1.s and 3.s.f forms instead of the 3.s.m form, as in (59a–b).

- (59) a. IL: \**huwa*            *ʔa-lʕab*                                *fī*                *ʔal*    gym  
           he                    Imperf.1.s-play                    in                the    gym  
 TL: *huwa*                *ya-lʕab*                                *fī*                *ʔan-nādī*  
           he                    Imperf.3.s.m-play                    in                the-gym  
           “He plays at the gym.”

(UIPG: 1st-year participant)

- b. IL: \**baʕd*                *huwa*                *ta-rūḥ*                                *ʔal-bayt*  
           after                    he                    Imperf.3.s.f-go                    the-house  
 TL: *baʕdēn*                *huwwa*                *bi-yi-ruḥ*                                *ʔil-bēt*  
           then                    he                    Hb-Imperf.3.s.m-go                    the-house  
           “Then, he goes home.”

(IPG: 2nd-year participant)







a. IL: \**hiya*                      *nawm*  
       she                                sleeping  
       TL: *hiya*                      *ta-nām*  
       she                                Imperf.3.s.m-sleep  
       “*She sleeps.*”

(UIPG: 2nd-year participant)

(65) a. IL: \**hiya*            *ya-ʔkul*            *fī*    *ʔiθnayn*            *fī*    *ʔal-masāʔ*  
       she                      Imperf.3.s.m-eat    in    two                    in    the-evening  
       TL: *hiya*            *ta-ʔkul*            *fī*    *ʔaθ-θāny-a*        *fī*    *ʔal-masāʔ*  
       she                      Imperf.3.s.f-eat    in    the-two-f            in    the-evening  
       “*She eats at two o’clock in the evening.*”

(IPG: 1st-year participant)

b. IL: \**wa*                      *ʔā-kul*                      pizza  
       and                            Imperf.1.s-eat            pizza  
       TL: *wa*                      *ta-ʔkul*                      pizza  
       and                            Imperf.3.s.f-eat            pizza  
       “*and she eats pizza.*”

(IPG: 3rd-year participant)

The third most frequent error was due to producing past verb forms instead of their present counterparts, as in (66a–b), and it constituted 8.5% of the total errors. Remarkably, this error was not produced at all by the first-year participants of both L2 groups. The fourth error pattern was due to inflecting the verbs for both the wrong tense and agreement, as in (67a–b), and it constituted only 7.6% of the total errors produced. The fifth error pattern constituted only 1.9% of the total incorrect tokens and was produced by two IPG participants who used the referent subject pronoun *huwa* “he” instead of *hiya* “she,” as in (68). The last error pattern was due to the unauthentic nature of some production tokens, and it constituted 1.4% of the total errors. It occurred when three second-year IPG participants produced EA-like verbs without the habitual prefix *bi-*, as in (69a–b).

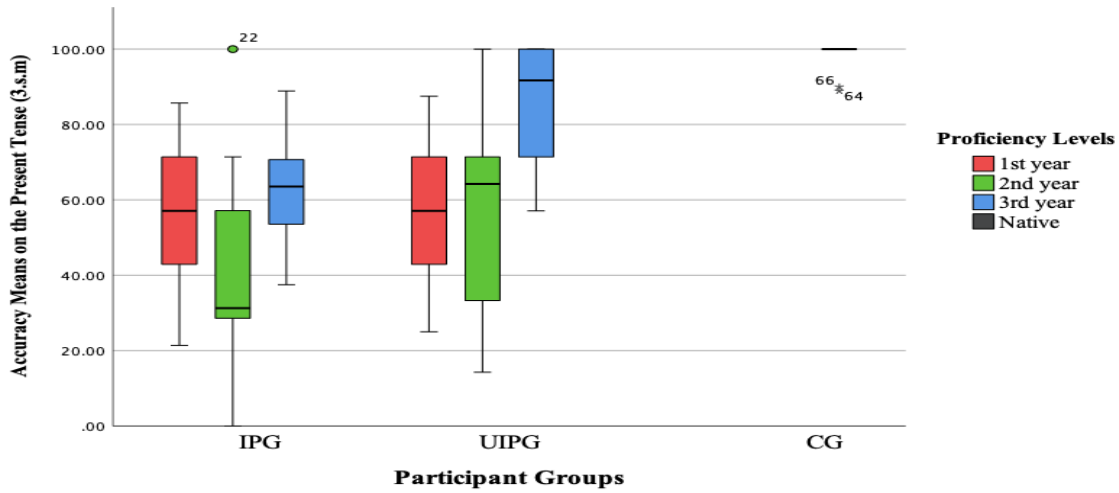


b. IL: *hiya	ta-rūḥ	ʔal	gym
she	Imperf.3.s.f-go	the	gym
TL: hiyya	bi-t-rūḥ	ʔan-nādī	
she	Hb-Imperf.3.s.f-go	the-gym	
“She goes [to] the gym			

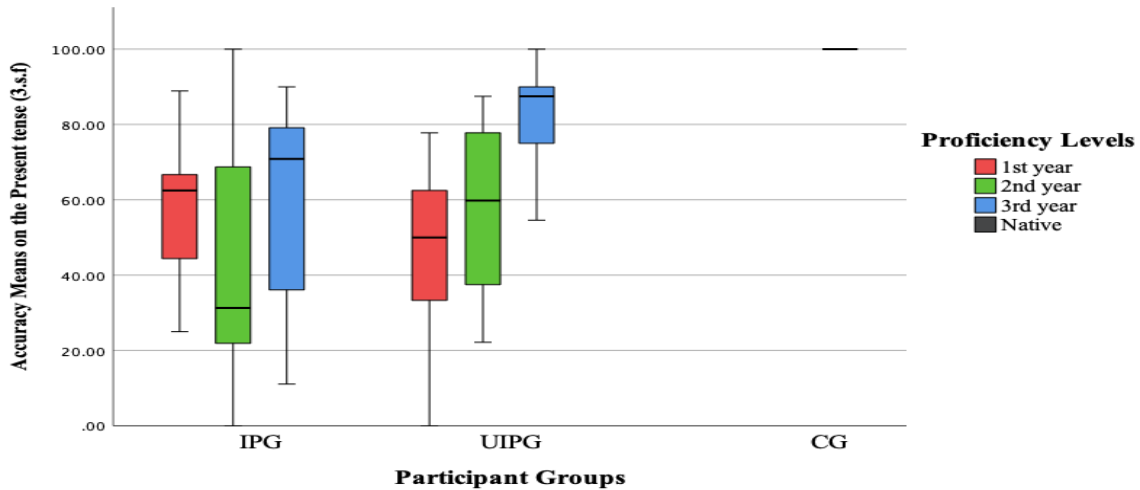
(IPG: 2nd-year participant)

The aggregate accuracy data of the L2 participants on the present tense exhibit different patterns than those revealed by the past tense data. Figures 4.5 and 4.6 show the trends (in quartiles) of the production accuracy of the IPG, UIPG, and CG on the present tense in the 3.s.m and 3.s.f contexts, respectively (see also Table 4.9 above). Generally, each proficiency level within each L2 group exhibited a relatively comparable accuracy rate in their production of the present tense with both agreement forms. In addition, as depicted in the two boxplots, the production accuracy of the IPG participants on this tense with both agreement forms did not improve considerably, particularly from the first-year level to the third-year level. In contrast, although the accuracy interquartile ranges of the UIPG participants were comparable at the first-year and second-year levels, their accuracy increased substantially at the third-year level, where their median scores were at or above 85% of accuracy in both agreement contexts. Furthermore, notwithstanding the comparable accuracy between the IPG and UIPG at the first-year and second-year levels on the present tense with both agreement forms, the third-year IPG participants exhibited lower accuracy, compared to their UIPG counterparts, in both agreement contexts. Moreover, it is visually evident that some participants of both L2 groups were 100% accurate in their production of this tense in both agreement contexts. This finding strongly suggests that ultimate attainment is possible in the acquisition of the present tense inflected for the 3.s.m and 3.s.f agreement forms.

**Figure 4.5** Boxplot for the Accuracy of the IPG, UIPG, and CG on the Present Tense with the 3.s.m Agreement Form



**Figure 4.6** Boxplot for the Accuracy of the IPG, UIPG, and CG on the Present Tense with the 3.s.f Agreement Form



Two full factorial two-way ANOVA tests were conducted to examine the main effects of the two independent variables (L2 participant group and L2 proficiency level), as well as their interaction effects, on the accuracy of the two L2 groups across the three proficiency levels on the present tense inflected for 3.s.m and 3.s.f. The first two-way ANOVA test that was



performed on the accuracy of the L2 participants on the present tense with the 3.s.m agreement form revealed a number of findings. The main effect of the L2 participant group factor was statistically significant ( $F(1,51) = 5.160, p = .027$ ) and yielded an effect size of .092, indicating that 9.2% of the variances in the L2 participants' accuracy were explained by the participant groups to which they belonged. The main effect of the proficiency level factor was likewise significant ( $F(2,51) = 6.390, p = .003$ ) and yielded an effect size of .200, indicating that 20% of the variances in the L2 participants' accuracy were explained by their proficiency levels. However, the interaction effect was not significant ( $F(2,51) = 1.309, p = .279, \text{partial } \eta^2 = .049$ ), indicating that there was no combined effect of the L2 participant group and L2 proficiency level factors on the L2 participants' production accuracy on the present tense with the 3.s.m agreement form. To examine the nature of the differences between all L2 proficiency levels, Tukey's post hoc analyses were conducted (Table 1h: Appendix H). The results revealed that the first-year and second-year participants did not differ significantly from one another ( $p = .673$ ). However, the third-year participants differed significantly from both the first-year participants ( $p = .024$ ) and second-year participants ( $p = .002$ ). Based on the follow-up pairwise comparisons between the IPG and UIPG at each proficiency level (Table 2h: Appendix H), the main effect of the L2 participant group factor was significant because of the outperformance of the third-year UIPG participants over their IPG counterparts ( $p = .034$ ). Additionally, based on the follow-up pairwise comparisons between all proficiency levels within each L2 group (Table 3h: Appendix H), the L2 proficiency level factor yielded a significant main effect because the third-year IPG participants significantly outperformed only the second-year IPG participants ( $p = .041$ ) and because the third-year UIPG participants significantly outperformed both the first-year UIPG participants ( $p = .004$ ) and second-year UIPG participants ( $p = .007$ ).

As for the second two-way ANOVA test that was performed on the L2 participants' accuracy on the present tense inflected for 3.s.f, it revealed several findings. The main effect of the L2 participant group factor was not significant ( $F(1,51) = 2.561, p = .116, \text{partial } \eta^2 = .048$ ). The main effect of the L2 proficiency level factor was significant ( $F(2,51) = 3.816, p = .029$ ) and yielded an effect size of .130, indicating that 13% of the variances in the L2 participants' accuracy were explained by their proficiency levels. However, the interaction effect was not significant ( $F(2,51) = 2.481, p = .094, \text{partial } \eta^2 = .089$ ), indicating that there was no combined effect of the L2 participant group and L2 proficiency level factors on the L2 participants' accuracy on the present tense with the 3.s.f form. Tukey's post hoc analyses (Table 4h: Appendix H) revealed that the first-year and second-year participants did not differ significantly from one another ( $p = .970$ ). However, the third-year participants differed almost significantly from the first-year participants ( $p = .051$ ) and significantly from the second-year participants ( $p = .026$ ). Based on the follow-up pairwise comparisons between the IPG and UIPG at each proficiency level (Table 5h: Appendix H), although the L2 participant group factor did not reveal a significant difference between the IPG and UIPG, the third-year UIPG participants significantly outperformed their IPG counterparts ( $p = .032$ ). In addition, based on the follow-up pairwise comparisons between all the proficiency levels within each L2 group (Table 6h: Appendix H), the L2 proficiency level factor yielded a significant main effect because the third-year UIPG participants significantly outperformed the first-year UIPG participants ( $p = .002$ ) and second-year UIPG participants ( $p = .023$ ), whereas the participants representing all proficiency levels within the IPG did not differ significantly from one another.

To examine whether the control participants differed from the L2 participants representing each proficiency level within each L2 group, 12 independent sample *t*-tests were

performed: six were conducted on the accuracy means of the present tense with the 3.s.m agreement form and the other six were conducted on the accuracy means of the present tense with the 3.s.f agreement form. The results revealed that, among all L2 participants, only the third-year UIPG participants were statistically comparable with the CG in their accuracy in the 3.s.m context (Table 4.12). However, both L2 groups across all proficiency levels were significantly less accurate than the CG in the 3.s.f context (Table 4.13).

**Table 4.12** Results of *t*-Tests on the Accuracy Mean Scores of the CG and Each Proficiency Level Within Each L2 Group on the Present Tense with the 3.s.m agreement form

Variable (I)	Variable (J)	df	t	Sig. p	95% Confidence Interval	
					Lower Bound	Upper Bound
CG	IPG					
Native (n = 9)	1st year (n = 9)*	8.906	-6.404	< .001	-58.08	-27.72
	2nd year (n = 11)*	10.633	-6.411	< .001	-76.26	-37.16
	3rd year (n = 8)*	8.094	-6.052	< .001	-48.18	-21.63
CG	UIPG					
Native (n = 9)	1st year (n = 10)*	10.069	-6.515	< .001	-57.15	-28.04
	2nd year (n = 10)*	9.583	-4.574	.001	-59.67	-20.42
	3rd year (n = 9)*	9.329	-2.019	.073	-23.87	1.29

Note. \* = equality of variances was not assumed via Levene's *F* test ( $p < .05$ ).

**Table 4.13** Results of *t*-Tests on the Accuracy Mean Scores of the CG and Each Proficiency Level Within Each L2 Group on the Present Tense with the 3.s.f agreement form

Variable (I)	Variable (J)	df	t	Sig. p	95% Confidence Interval	
					Lower Bound	Upper Bound
CG	IPG					
Native (n = 9)	1st year (n = 9)*	8.000	-6.720	< .001	-58.13	-28.43
	2nd year (n = 11)*	10.000	-5.738	< .001	-79.05	-34.83
	3rd year (n = 8)*	7.000	-4.243	.004	-66.03	-18.77
CG	UIPG					
Native (n = 9)	1st year (n = 10)*	9.000	-6.925	< .001	-69.84	-35.44
	2nd year (n = 10)*	9.000	-5.508	< .001	-59.95	-25.05
	3rd year (n = 9)*	8.000	-3.695	.006	-25.68	-5.94

Note. \* = equality of variances was not assumed via Levene's *F* test ( $p < .05$ ).

Taken altogether, the following main findings were revealed by the qualitative and quantitative analyses. Initially, the IPG produced 12 tokens only in the EA variety. Only one token produced by a second-year participant was completely correct, whereas the remaining 11 tokens were partially incorrect due to lack of suppliance of the habitual prefix *bi-* in its obligatory used context. Remarkably, none of these tokens were produced by the third-year IPG participants. This finding strongly suggests that the IPG participants exhibited a tendency to adhere to the MSA variety, unlike the CG, whose present tense data were exclusively and entirely in the EA variety. Additionally, the most prevalent pattern of errors by both L2 groups (i.e., the IPG and UIPG) when producing the present tense structure in both the 3.s.m and 3.s.f contexts was due to lack of Arabic verb suppliance in obligatory contexts. However, this pattern gradually decreased as proficiency increased within the UIPG, whereas it considerably increased at the second-year level before it decreased again at the third-year level within the IPG. Furthermore, both L2 groups exhibited a tendency to converge in their use of the default form (3.s.m) in the 3.s.f context. However, while this agreement error decreased as proficiency increased within the UIPG, it was prevalent in the data of the IPG participants and did not exhibit a decreasing tendency as their proficiency increased.

Moreover, despite the very small amount of EA verbs produced by the IPG participants, they significantly underperformed their UIPG counterparts in the long run, particularly at the third-year level, in their accuracy on the present tense with both agreement forms. Moreover, the third-year UIPG participants significantly outperformed both the first-year and second-year UIPG participants in both agreement contexts, unlike the third-year IPG participants, who did not significantly differ from the first-year IPG participants in the 3.s.m context and did not differ significantly from both the first-year and second-year IPG participants in the 3.s.f context.

Lastly, similar to what was revealed by the acquisition data of the past tense, among all participants across all proficiency levels, only the third-year UIPG participants were statistically on par with the CG in their accuracy on the present tense in the 3.s.m context. However, both the IPG and UIPG, across all proficiency levels, significantly underperformed the CG in the 3.s.f context. This finding suggests that, when it comes to the acquisition of the present tense structure, only the UIPG were able to gradually approach native-like proficiency over time (in the 3.s.m context only).

### 4.3.2 Negation

#### 4.3.2.1 Past negation

Based on the coding criteria discussed in Section 3.6.2.3, the past negation data were analyzed accordingly. Table 4.14 provides descriptive statistics for the production accuracy of the IPG, UIPG, and CG on the past negation.

**Table 4.14** Descriptive Statistics for the Accuracy of the IPG, UIPG, and CG on the Past Negation

Participant Group Proficiency Level	Past Negation			
	C/T	%	<i>M</i>	<i>SD</i>
<b>IPG</b>				
1st year ( <i>n</i> = 9)	5/90	5.6	5.56	13.33
2nd year ( <i>n</i> = 11)	18/111	16.2	15.45	33.28
3rd year ( <i>n</i> = 8)	12/82	14.6	14.79	16.99
<b>UIPG</b>				
1st year ( <i>n</i> = 10)	45/101	44.6	44.09	33.21
2nd year ( <i>n</i> = 10)	31/102	30.4	31.00	35.73
3rd year ( <i>n</i> = 9)	44/91	48.4	48.82	42.20
<b>CG</b>				
Native ( <i>n</i> = 9)	98/98	100	100	0.00

*Note.* C/T = Correct/Total; % of correct tokens; M = Mean; SD = Standard Deviation.

Recall, MSA exhibits two distinct constructions to negate the past tense: *mā + perfect verb* and *lam + imperfect verb*. Contrary to the former, the latter is less marked, more frequently recycled in the L2ers’s instructional textbooks, and more similar to the negation construction exhibited by their L1 (English). Thus, to examine the role of input frequency and L1 transfer in the acquisition of these two constructions, it was essential to analyze the production data of both constructions. Remarkably, the L2 participants supplied the more marked and less frequent construction (i.e., *mā + perfect verb*) more often than the less marked and more frequent construction (i.e., *lam + imperfect verb*). Precisely, out of the 153 completely correct tokens produced, only 16 correct tokens exhibited the *lam + imperfect verb* construction, as in (70a–b), whereas the remaining 137 tokens exhibited the *mā + perfect verb* construction, as in (71a–d). In fact, the 16 tokens exhibiting the *lam + imperfect verb* construction were produced by five participants only: six tokens by one first-year IPG participant, four tokens by two third-year IPG participants, and six tokens by two third-year UIPG participants.

(70) a. IL: *ʔar-rajul lam ya-ḏhab ʔilā ʔal park*  
the-man did not Imperf.3.s.m-go to the park  
TL: *ʔar-rajul lam ya-ḏhab ʔilā ʔal-ḥadīq-a*  
the-man did not Imperf.3.s.m-go to the-park-s.f  
“The man did not go to the park.”  
(IPG: 3rd-year participant)

b. *hiya lam ta-šrab ʔal-qahw-a*  
she did not Imperf.3.s.f-drink the-coffee-f  
“She did not drink coffee.”  
(UIPG: 3rd-year participant)

(71) a. *huwa mā laʕiba kur-at ʔal-qadam*  
he did not Perf.played.3.s.m ball-s.f the-foot.s.f  
“He did not play soccer.”  
(IPG: 1st-year participant)

- b. *hiya mā šarib-at ʔaš-šāy*  
 she did not Perf.drunk-3.s.f the-tea  
 “She did not drink tea.”  
 (UIPG: 1st-year participant)

- c. *mā ḏahab-at ʔilā ʔal-baħr*  
 did not Perf.went-3.s.f to the-sea  
 “She did not go to the sea.”  
 (IPG: 2nd-year participant)

- d. *huwa mā faʕala ʔayy šayʔ*  
 he did not Perf.did.3.s.m any thing  
 “He did not do anything.”  
 (UIPG: 2nd-year participant)

Moreover, as discussed in Section 3.3.2.1 above, EA exhibits the discontinuous negator *ma-š* that is affixed to the perfect verb to negate past sentences. Remarkably, out of the 283 tokens produced by the IPG participants, only 58 tokens were in the EA variety (20.5%): 11 by first-year participants, 37 by second-year participants, and 10 by third-year participants. However, only two of these EA tokens were correct and produced by a third-year participant in the 3.s.m context, as in (72a–b). Regarding the data of the control participants, all data were produced in the EA variety (i.e., the *mā + perfect verb + š* construction), as in (73a–b).

- (72) a. IL: *huwa mā-faʕil-š hāḏā*  
 he Neg-Perf.did.3.s.m-Neg this  
 TL: *huwwa ma-ʕmil-š dah*  
 he Neg-Perf.did.3.s.m-Neg this  
 “He did not do this.”  
 (IPG: 3rd-year participant)

- b. *wa baʕda ḏālika ma-rāħi-š ʔal-ħadīq-a*  
 and after that Neg-Perf.went.3.s.m-Neg the-park-s.f  
 “And then, he did not go to the park.”  
 (IPG: 3rd-year participant)

- (73) a. *ʔir-rāgil*                      *ma-ʕimil-š*                      *ḥāg-a*  
 The-man                      Neg-Perf.did.3.s.m-Neg                      thing-s.f  
 “The man did not do anything.”  
 (control participant)
- b. *ma-raḥi-t-š*    *ʔiš-šatt*  
 Neg-Perf.went-3.s.f-Neg    the-beach  
 “She did not go to the beach.”  
 (control participant)

The L2 participants’ incorrect production tokens in the past negation structure were categorized into 12 error patterns, as summarized in Table 4.15. The production of the wrong negator in MSA sentences was the predominant error across all participants and constituted 29.6% of the total errors. This error occurred when the participants produced mainly the verbal present negator *lā* “does/do not” instead of its past counterparts, as in (74a–b).

**Table 4.15** *Error Patterns Produced by the Participants in the Past Negation (Error Type Tokens/Total Errors)*

Error Type	IPG			UIPG		
	1st year (n = 9)	2nd year (n = 11)	3rd year (n = 8)	1st year (n = 10)	2nd year (n = 10)	3rd year (n = 9)
Wrong N in MSA	36/85	13/93	25/70	7/56	14/71	30/47
Wrong N + V in MSA	11/85	23/93	18/70	19/56	17/71	6/47
L of Neg S	16/85	7/93	5/70	7/56	8/71	2/47
Wrong N + V in EA	6/85	37/93	0/70	0/56	0/71	0/47
Wrong T in MSA	1/85	2/93	8/70	5/56	14/71	5/47
Wrong Agr in MSA	8/85	3/93	1/70	8/56	6/71	3/47
Wrong T + Agr in MSA	1/85	5/93	2/70	2/56	8/71	0/47
Lack of MSA-V	0/85	3/93	2/70	8/56	4/71	1/47
Wrong T in EA	2/85	0/93	5/70	0/56	0/71	0/47
Wrong Agr in EA	1/85	0/93	3/70	0/56	0/71	0/47
Wrong T + Agr in EA	2/85	0/93	1/70	0/56	0/71	0/47
Lack of EA-V	1/85	0/93	0/70	0/56	0/71	0/47

*Note.* L of Neg S = Lack of Negation Suppliance; N = Negator; V = Verb; T = Tense; Agr = Agreement; N + V = Negator was incorrect and Verb was inflected for either the wrong tense or agreement.







- b. IL: *\*lam*            *laʕab-at*            tennis  
       did not            Perf.played-3.s.f            tennis  
 TL: *lam*                *ta-lʕab*                tennis  
       did not            Imperf.3.s.f-play            tennis  
       “She did not play tennis.”

(IPG: 3rd-year participant)

The sixth most frequent error was due to the correct production of MSA negators with verbs inflected for the wrong agreement, and it constituted 6.9% of the total errors. This error occurred mostly in the 3.s.f context, as in (78a–b).

- (78) a. IL: *\*mā*            *laʕab-ti*                *kur-at*                *ʔal-qadam*  
       did not            Perf.played-2.s.f            ball-s.f                the-foot.s.f  
 TL: *mā*                *laʕiba*                *kur-at*                *ʔal-qadam*  
       did not            Perf.went.3.s.m            ball-s.f                the-foot.s.f  
       “He did not play soccer.”

(IPG: 1st-year participant)

- b. IL: *\*mā*                *sabaḥ-ta*  
       did not            Perf.swam-2.s.m  
 TL: *mā*                *sabaḥ-at*  
       did not            Perf.swam-3.s.f  
       “She did not swim.”

(UIPG: 1st-year participant)

Two other types of errors were due the correct production of MSA negators with verbs inflected for both the wrong tense and agreement, as in (79a–b), and lack of MSA verb suppliance, as in (80a–b). Each type constituted 4.3% of the total errors.

- (79) a. IL: *\*lam*            *sāfar*  
       did not            Perf.traveled.3.s.m  
 TL: *lam*                *tu-sāfir*  
       did not            Imperf.3.s.f-travel  
       “She did not travel.”

(IPG: 2nd-year participant)

- b. IL: \**mā*            *ʔa-lʕab*                            *mubār-ā*  
           did not            Imperf.1.s-play                            game-s.f  
 TL: *mā*                *laʕiba*                                        *mubār-ā*  
           did not            Perf.played.3.s.m                            game-s.f  
 “He did not play a game.”  
 (UIPG: 2nd-year participant)

- (80) a. IL: \**lam*            *ʕinda-hā*                                        *šay*  
           did not            with-her                                        thing  
 TL: *lam*                *ta-ʕal*                                        *ʔayy*    *šayʔ*  
           did not            Imperf.3.s.f-do                                any        thing  
 “She did not do anything.”  
 (IPG: 3rd-year participant)

- b. IL: \**ʔar-rajul*    *mā*    biking  
           the-man            did not    biking  
 TL: *ʔar-rajul*    *mā*    *rakiba*    *ʔad-darrāj-a*  
           The-man            did not    Perf.rode.3.s.m                                the-bike-f  
 “The man did not bike.”  
 (UIPG: 3rd-year participant)

The last patterns of errors were produced by the IPG participants only and were due to the correct production of the EA past negator *ma-š* with verbs inflected for the wrong tense (1.7%), wrong agreement (0.9%), both wrong tense and agreement (0.7%), and lack of EA verb suppliance (0.2%), as in (81), (82), (83), and (84), respectively.

- (81) IL: \**ma-ti-šrabu-š*    *ʔal-qahw-a*  
           Neg-Imperf.3.s.f-drink-Neg                                        the-coffee-f  
 TL: *ma-širb-it-š*    *ʔil-ʔahw-a*  
           Neg-Perf.drunk-3.s.f-Neg                                        the-coffee-f  
 “She did not drink coffee.”  
 (IPG: 3rd-year participant)

- (82) IL: \**hiya*            *ma-faʕal-š*  
           she            Neg-Perf.did.3.s.m-Neg  
 TL: *hiyya*            *ma-ʕmil-it-š*                                        *ḥāg-a*  
           she            Neg-Perf.did-3.s.f-Neg                                thing-s.f  
 “She did not do anything.”  
 (IPG: 1st-year participant)

- (83) IL: \**baʕda*    *ðālika*    *hiya*    *ma-ʔa-šrub-š*    *ʔal-qahw-a*  
 After    that    she    Neg-Imperf.1.s-drink-Neg    the-coffee-f  
 TL: *baʕdēn*    *hiyya*    *ma-širb-it-š*    *ʔil-ʔahw-a*  
 then    she    Neg-Perf.drunk-3.s.f-Neg    the-coffee-f  
 “Then, she did not drink coffee.”

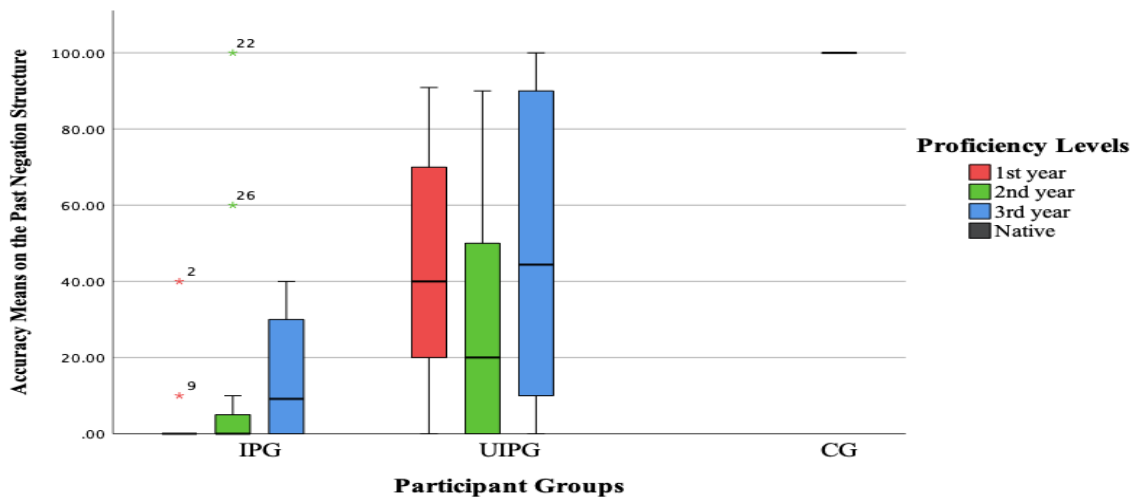
(IPG: 3rd-year participant)

- (84) IL: \**hiya*    *ma-ʕinda-hā-š*    *kura*    *wa*    *kursī*  
 she    Neg-with-3.s.f-Neg    ball    and    chair  
 TL: *hiyya*    *ma-kan-š*    *ʕinda-hā*    *kūra*    *wi*    *kursī*  
 she    Neg-Perf.was.3.s.m-Neg    with-her    ball    and    chair  
 “She did not have a ball and chair.”

(IPG: 1st-year participant)

The aggregate accuracy data of the L2 participants on the past negation show again an advantage for the UIPG over the IPG. Figure 4.7 depicts the trends (in quartiles) of the accuracy of the IPG, UIPG, and CG on the past negation (see also Table 4.14 above). Aside from the CG, whose accuracy rate was at the ceiling level, the UIPG exhibited a higher accuracy rate than the IPG across all proficiency levels. Interestingly, the accuracy interquartile ranges of the first-year and second-year UIPG participants were even higher than that of the third-year IPG participants. These two observations suggest that the UIPG participants acquired the past negation more readily and accurately than their IPG counterparts. Additionally, when it comes to the development of this structure within each L2 group, the performance of the IPG displayed very marginal improvement from the first-year through the third-year level. In contrast, the UIPG participants exhibited a U-shaped curve of accuracy, where they initially exhibited high accuracy at the first-year level, experienced a backslid at the second-year level, and eventually improved at the third-year level. Moreover, attainment of native-like proficiency was confirmed since the accuracy of one second-year IPG participant and two third-year UIPG participants was at the ceiling level (100% of accuracy).

**Figure 4.7** Boxplot for the Accuracy of the IPG, UIPG, and CG on the Past Negation



A full factorial two-way ANOVA test was conducted to examine the main effects of the two independent variables (the L2 participant group and L2 proficiency level), as well as their interaction effect, on the accuracy of the two L2 groups across the three proficiency levels on the past negation. The test revealed a number of findings. The main effect of the L2 participant group factor was statistically significant ( $F(1,51) = 12.368, p < .001$ ) and yielded an effect size of .195, indicating that 19.5% of the variances in the L2 participants' accuracy were explained by the L2 participant groups to which they belonged. However, the main effect of the L2 proficiency level factor was not significant ( $F(2,51) = .382, p = .684, \text{partial } \eta^2 = .015$ ). The interaction effect was also not significant ( $F(2,51) = .755, p = .475, \text{partial } \eta^2 = .029$ ), indicating that there was no combined effect of the L2 participant groups and L2 proficiency levels on the L2 participants' accuracy on the past negation. In other words, based on the follow-up pairwise comparisons between the IPG and UIPG at each proficiency level (Table 1i: Appendix I), the L2 participant factor yielded a significant main effect because the UIPG significantly outperformed the IPG at the first-year level ( $p = .010$ ) and third-year level ( $p = .030$ ).

A further visual inspection of the boxplot revealed that the maximum accuracy score reached by the third-year IPG participants intersects with the accuracy median of the first-year UIPG participants. This observation promoted conducting an independent samples *t*-test on the accuracy mean scores of these specific two groups of participants. Remarkably, the test yielded a significant difference between the means of these two groups ( $t(16) = 2.261, p = .038$ ). In other words, the first-year UIPG participants were significantly more accurate than the third-year IPG participants in their accuracy on the past negation.

To examine whether the control participants differed from the L2 participants of each proficiency level within each L2 group in their accuracy on the past negation, six independent sample *t*-tests were performed. The results revealed that the accuracy mean of the CG was significantly higher than those of the IPG and UIPG across all proficiency levels (Table 4.16).

**Table 4.16** Results of *t*-Tests on the Accuracy Mean Scores of the CG and Each Proficiency Level Within Each L2 Group on the past negation

Variable (I)	Variable (J)	df	t	Sig. p	95% Confidence Interval	
					Lower Bound	Upper Bound
CG	IPG					
Native (n = 9)	1st year (n = 9)*	8.000	-21.250	< .001	-104.69	-84.20
	2nd year (n = 11)*	10.000	-8.427	< .001	-106.90	-62.19
	3rd year (n = 8)*	7.000	-14.190	< .001	-99.41	-71.01
CG	UIPG					
Native (n = 9)	1st year (n = 10)*	9.000	-5.324	< .001	-79.66	-32.16
	2nd year (n = 10)*	9.000	-6.107	< .001	-94.56	-43.44
	3rd year (n = 9)*	8.000	-3.639	.007	-83.61	-18.74

Note. \* = equality of variances was not assumed via Levene's *F* test ( $p < .05$ ).

To sum up, the qualitative and quantitative analyses of the past negation data yielded a number of main findings. First, the participants of both the IPG and UIPG across all proficiency levels exhibited a tendency to use the *mā + perfect verb* construction more often than the *lam +*

*imperfect verb* construction. Second, the IPG participants opted for the MSA system more often than its EA counterpart. Third, the UIPG participants were significantly more accurate than the IPG at the first-year and third-year levels. Fourth, no statistically significant differences were found between all proficiency levels within each L2 groups. Finally, notwithstanding the significant outperformance of the UIPG participants over their IPG counterparts, the CG significantly outperformed both L2 groups across all proficiency levels.

#### 4.3.2.2 Verbal present negation

Based on the coding criteria discussed in Section 3.6.2.4, the production data of the verbal present negation were analyzed accordingly. Table 4.17 provides descriptive statistics for the accuracy of the IPG, UIPG, and CG on the verbal present negation.

**Table 4.17** Descriptive Statistics for the Accuracy of the IPG, UIPG, and CG on the Verbal Present Negation

Participant Group Proficiency Level	Verbal Present Negation			
	C/T	%	<i>M</i>	<i>SD</i>
<b>IPG</b>				
1st year ( <i>n</i> = 9)	47.5/83	57.2	58.28	30.80
2nd year ( <i>n</i> = 11)	40.5/106	38.2	37.07	40.01
3rd year ( <i>n</i> = 8)	52/73	71.2	70.83	32.51
<b>UIPG</b>				
1st year ( <i>n</i> = 10)	60/91	65.9	65.68	18.65
2nd year ( <i>n</i> = 10)	63/94	67	67.14	20.80
3rd year ( <i>n</i> = 9)	79/82	96.3	96.30	7.85
<b>CG</b>				
Native ( <i>n</i> = 9)	82/82	100	100	00

*Note.* C/T = Correct/Total; % of correct tokens; M = Mean; SD = Standard Deviation.

Examples of the L2 participants' correctly produced tokens of the verbal present negation are provided in (85a–d). Recall, in addition to the MSA verbal present negation system (i.e., the *lā* + *imperfect verb* construction), the IPG participants were introduced to the EA verbal present



negation system (i.e., the *mā + bi + imperfect verb + š* construction). Out of the 262 tokens produced by the IPG participants, 46 tokens were in the EA variety (17.6%): 10 by first-year participants, 32 by second-year participants, and 4 by a third-year participant. Remarkably, only 2 of these EA tokens were fully correct and were produced by a first-year participant with the verb *yihibb* “to like” in the 3.s.m context, as in (86a–b). With regard to the data of the control participants, all verbal present negation tokens were produced in the EA variety, as in (87a–b).

- (85) a. *wa hiya lā tu-šāhid-u ʔat-tilifīzyūn*  
 and she does not Imperf.3.s.f-watch-Indic the-television  
 “And she does not watch TV.”  
 (UIPG: 1st-year participant)
- b. *fī ʔas-sāʕ-a ʔas-sādis-a huwa lā ya-nām*  
 in the-clock-f the-six-f he does not Imperf.3.s.m-sleep  
 “At six o’clock, he does not sleep.”  
 (IPG: 2nd-year participant)
- c. *fī ʔat-tāsiʕ-a šabāḥ-an lā yu-daxxin ʔar-rajul*  
 in the-nine-f morning-Acc does not Imperf.3.s.m-smoke the-man  
 “At nine in the morning, the man does not smoke.”  
 (UIPG: 3rd-year participant)
- d. *wa ʔas-sāʕ-a sabʕ-a hiya lā ta-fʕal ʔayy ḥāq-a*  
 and the-clock-f seven-f she does not Imperf.3.s.f-do any thing-s.f  
 “And at seven o’clock, she does not do anything.”  
 (IPG: 3rd-year participant)
- (86) a. IL: *mā-bi-yi-ḥibbi-š* *qahw-a*  
 Neg-Hb-Imperf.3.s.m-like-Neg coffee-f  
 TL: *ma-bi-yi-ḥibbi-š* *ʔil-ʔahw-a*  
 Neg-Hb-Imperf.3.s.m-like-Neg the-coffee-f  
 “He does not like coffee.”  
 (IPG: 1st-year participant)

- b. IL: *mā-bi-yi-ḥibbi-š* smoking  
 Neg-Hb-Imperf.3.s.m-like-Neg smoking  
*ma-bi-yi-ḥibbi-š* *ʔit-tadxīn*  
 Neg-Hb-Imperf.3.s.m-like-Neg the-smoking.VN  
 “He does not like smoking.”

(IPG: 1st-year participant)

- (87) a. *wi ʔis-sāʕ-a tamany-a ma-bi-yi-ʔrā-š*  
 and the-clock-f eight-f Neg-Hb-Imperf.3.s.m-read-Neg  
 “And at eight o’clock, he does not read.”

(control participant)

- b. *ʔis-sāʕ-a tamany-a ma-bi-ti-tfarrag-š ʕalā ʔit-tilfīzyūn*  
 the-clock-f eight-f Neg-Hb-Imperf.3.s.f-watch-Neg on the-television  
 “At eight o’clock, she does not watch TV.”

(control participant)

Overall, 12 patterns of errors were detected in the verbal present negation data produced by the L2 participants. Table 4.18 provides a snapshot of the error patterns detected. The most frequent type of errors was due to lack of MSA verb suppliance with the correctly produced MSA negator *lā* “does not.” This type of error constituted 20.1% of the total errors, and it occurred when the L2 participants produced English verbs instead of their Arabic equivalents, as in (88a), and when they did not produce a verb at all, as in (88b). The second most frequent error was due to inflecting the negated verbs for the wrong agreement in MSA-like negation sentences. This type of error constituted 17% of the total errors, and it occurred when the verbs were conjugated for the 1.s form instead of its 3.s.m and 3.s.f counterparts, as (89a–b), and when the verbs were conjugated for the 3.s.m form in the 3.s.f context, as in (89c–d). The third most frequent error was due to lack of negation suppliance, constituting 11.9% of the total errors. It occurred mostly when the participants failed to describe negation pictures or skipped them altogether.

**Table 4.18** Error Patterns Produced by the Participants in the Verbal Present Negation (Error Type Tokens/Total Errors)

Error Type	IPG			UIPG		
	1st year (n = 9)	2nd year (n = 11)	3rd year (n = 8)	1st year (n = 10)	2nd year (n = 10)	3rd year (n = 9)
L of MSA-V	9/39	5/67	0/23	15/31	7/31	3/3
Wrong Agr in MSA	5/39	9/67	6/23	6/31	7/31	0/3
L of Neg S	6/39	7/67	2/23	3/31	5/31	0/3
Wrong N in MSA	7/39	4/67	5/23	0/31	1/31	0/3
Wrong N + V in EA	0/39	17/67	0/23	0/31	0/31	0/3
Wrong T in MSA	2/39	2/67	1/23	3/31	5/31	0/3
Wrong N + bi in EA	0/39	12/67	0/23	0/31	0/31	0/3
Wrong N + V in MSA	0/39	5/67	0/23	4/31	3/31	0/3
Authenticity*	7/39	3/67	4/23	0/31	0/31	0/3
Hybridized N	0/39	3/67	4/23	0/31	0/31	0/3
Wrong T + Agr in MSA	2/39	0/67	1/23	0/31	3/31	0/3
Wrong T + bi in EA	1/39	0/67	0/23	0/31	0/31	0/3

*Note.* \* = coded as partially incorrect; L of Neg S = Lack of Negation Suppliance, N = Negator; V = Verb; T = Tense; Agr = Agreement; N+V = Negator was incorrect and Verb was inflected for either the wrong tense or agreement.

- (88) a. IL: \**huwa lā* read *ʔal-kitāb*  
 he does not read the-book.s.m  
 TL: *huwa lā ya-qraʔ* *ʔal-kitāb*  
 he does not Imperf.3.s.m-read the-book.s.m  
 “He does not read the book.”

(UIPG: 1st-year participant)

- b. IL: \**hiya lā* *ḥāq-a*  
 she does not thing-s.f  
 TL: *hiya lā ta-ʔʔal* *ʔayy šayʔ*  
 then does not Imperf.3.s.f-eat any thing.s.m  
 “She does not do anything.”

(IPG: 1st-year participant)

- (89) a. IL: \**huwa lā* *ʔa-šrab* *ʔal-ʔahw-a*  
 he does not Imperf.1.s-drink the-coffee-f  
 TL: *huwa lā ya-šrab* *ʔal-qahw-a*  
 he does not Imperf.3.s.m-drink the-coffee-f  
 “He does not drink coffee.”

(IPG: 2nd-year participant)



- (90) a. IL: \**huwa mā ya-tanāwal fī ʔaḏ-ḏuhr yadāʔ*  
 he did not Imperf.3.s.m-have in the-noon lunch  
 TL: *huwa lā ya-tanāwal yadāʔ fī ʔaḏ-ḏuhr*  
 he does not Imperf.3.s.m-have lunch in the-noon  
 “He does not have lunch at noon.”  
 (UIPG: 2nd-year participant)
- b. IL: \**lam ta-šrab šāy*  
 did not Imperf.3.s.f-drink tea  
 TL: *lā ta-šrab ʔaš-šāy*  
 did not Imperf.3.s.f-drink the-tea  
 “She does not drink tea.”  
 (IPG: 2nd-year participant)
- c. IL: \**huwa layasa yu-daxxin ʔas-sijār-a*  
 he is not.3.s.m Imperf.3.s.m-smoke the-cigarette-s.f  
 TL: *huwa lā yu-daxxin ʔas-sajāʔir*  
 he does not Imperf.3.s.m-smoke the-cigarettes.Pl  
 “He does not smoke cigarettes.”  
 (UIPG: 3rd-year participant)
- (91) a. IL: \**fī ʔat-tisʕ-a muš ʔa-šrab ʔas-sijār-a*  
 in the-nine-f Neg Imperf.1.s-drink the-cigarette-s.f  
 TL: *wi tisʕ-a ma-bi-yi-šrab-š sagāyir*  
 and nine-f Neg-Hb-Imperf.3.s.m-drink-Neg cigarettes.Pl  
 “At nine, he does not smoke cigarettes.”  
 (IPG: 2nd-year participant)
- b. IL: \**hiya miš ʔā-kul ʔal-ḩašāʔ*  
 she Neg Imperf.1.s-eat the-dinner  
 TL: *hiyya ma-b-tā-kul-š ʔil-ḩaša*  
 she Neg-Hb-Imperf.3.s.f-eat-Neg the-dinner  
 “She does not eat dinner.”  
 (IPG: 2nd-year participant)

The sixth error pattern was due to the previously mentioned cause but with correctly inflected verbs that were not preceded by the habitual prefix *bi-*, as in (92a–b). These two error patterns constituted 8.8% and 6.2% of the total errors, respectively, and were exclusively observed in the data of second-year IPG participants. Recall, the non-discontinuous EA negator *miš* is used to





a. IL: * <i>wa</i>	<i>baʕda</i>	<i>ðālika</i>	<i>ma-yi-daxxil-š</i>
and	after	that	Neg-Imperf.3.s.m-[smoke]-Neg
TL: <i>wi</i>	<i>baʕdēn</i>		<i>ma-bi-yi-daxxan-š</i>
and	then		Neg-Hb-Imperf.3.s.m-smoke-Neg
“And then, he does not smoke.”			

(IPG: 3rd-year participant)

c. IL: * <i>lākin</i>	<i>hiya</i>	<i>lā</i>	<i>ta-šuf</i>	<i>ʔat-tilifizyūn</i>
but	she	does not	Imperf.3.s.f-watch	the-television
TL: <i>lākin</i>	<i>hiya</i>	<i>lā</i>	<i>tu-šāhid</i>	<i>ʔat-tilifizyūn</i>
but	she	does not	Imperf.3.s.f-watch	the-television
“But she does not watch TV.”				

(IPG: 1st-year participant)

The tenth frequent type of errors was due to producing a hybridized negator—that is, using the MSA negator *lā* with the enclitic *-š* of the discontinuous EA negator *ma-š* or with the EA non-discontinuous negator *miš*, as in (96a–b), respectively. This type of error constituted 3.6% of the total errors and was observed in the data of some second-year and third-year IPG participants. The eleventh most frequent error constituted 3.1% of the total errors and was due to inflecting the verbs for both the wrong tense and agreement, as in (97). The least frequent error was due to inflecting the verb for the wrong tense in an EA-like negation sentence, provided in (98). It was produced only once (0.5%) by a first-year IPG participant.

(96) a. IL: * <i>baʕda</i>	<i>ðālika</i>	<i>lā</i>	<i>ya-šrab-š</i>	<i>ʔal-qahw-a</i>
after	that	does not	Imperf.3.s.m-drink-Neg	the-coffee-f
TL: <i>baʕda</i>	<i>ðālika</i>	<i>lā</i>	<i>ya-šrab</i>	<i>ʔal-qahw-a</i>
after	that	does not	Imperf.3.s.m-drink	the-coffee-f
“Then, he does not drink coffee.”				

(IPG: 3rd-year participant)

b. IL: * <i>wa</i>	<i>baʕdi</i>	<i>muš</i>	<i>yā-kul-š</i>	<i>ʔal</i> dinner
and	[then]	Neg	Imperf.3.s.m-eat-Neg	the dinner
TL: * <i>wi</i>	<i>baʕdēn</i>		<i>ma-bi-yā-kul-š</i>	<i>ʔil-ʕaša</i>
and	then		Neg-Hb-Imperf.3.s.m-eat-Neg	the-dinner
“And then, he does not eat dinner.”				

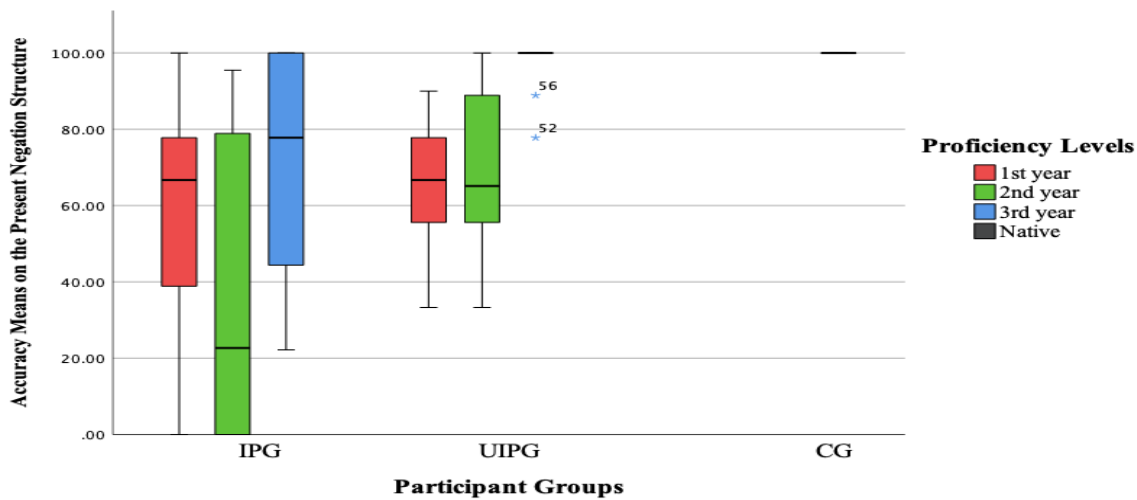
(IPG: 2nd-year participant)



- (97) IL: \**hiya*      *lā*      *faʕal*      *kull*      *šay*  
          she      does not      Perf.did.3.s.m      every      thing.s.m  
 TL: *hiya*      *lā*      *ta-ʕal*      *ʔayy*      *šayʔ*  
          she      does not      Imperf.3.s.f-do      any      thing.s.m  
 “She does not do anything.”  
 (UIPG: 2nd-year participant)
- (98) IL: \**ʔas-sāʕ-a*      *sabʕ-a*      *huwa*      *ma-faʕl-š*  
          the-clock-f      seven-f      he      Neg-Perf.did.3.s.m.Neg  
 TL: *ʔis-sāʕ-a*      *sabʕ-a*      *huwwa*      *ma-bi-yi-ʕmil-š*      *hāq-a*  
          the-clock-f      seven-f      he      Neg-Hb-Imerf.3.s.m-does.Neg      thing-s.f  
 “At seven o’clock, he does not do anything.”  
 (IPG: 1st-year participant)

The aggregate accuracy of the verbal present negation data exhibits configurations somehow different from those revealed by the past negation data. Figure 4.8 depicts the trends (in quartiles) of the production accuracy of the IPG, UIPG, and CG on the verbal present negation (see also Table 4.17 above). Both L2 groups demonstrated a higher accuracy rate across all proficiency levels. Precisely, the first-year and second-year UIPG participants exhibited a comparable performance, as indicated by the overlap in the medians of their accuracy interquartile ranges. Conversely, despite the overlap in the middle 50% of the accuracy mean scores of the first-year and second-year IPG participants, the interquartile median of the latter participants is visually lower than that of the former participants. However, the accuracy considerably increased at the third-year level within both L2 groups. It is worth noting that many participants across different proficiency levels in both L2 groups achieved accuracy at the ceiling level, particularly the third-year UIPG participants, most of whom were 100% accurate when producing the verbal present negation.

**Figure 4.8** Boxplot for the Accuracy of the IPG, UIPG, and CG on the Verbal Present Negation



A full factorial two-way ANOVA test was conducted to examine the main effects of the two independent variables (the L2 participant group and L2 proficiency level), as well as their interaction effect, on the accuracy of the two L2 groups across the three proficiency levels on the verbal present negation. The test revealed a number of findings. The main effect of the L2 participant group factor was statistically significant ( $F(1,51) = 8.221, p = .006$ ) and yielded an effect size of .139, indicating that 13.9% of the variances in the L2 participants' accuracy were explained by the L2 groups to which they belonged. The main effect of the proficiency level factor was likewise significant ( $F(2,51) = 6.284, p = .004$ ) and yielded an effect size of .198, indicating that 19.8% of the variances in the L2 participants' accuracy were explained by their proficiency levels. However, the interaction effect was not significant ( $F(2,51) = .921, p = .405$ , partial  $\eta^2 = .035$ ), indicating that there was no combined effect of the L2 participant group and L2 proficiency level factors on the L2 participants' accuracy on the verbal present negation. Tukey's post hoc analyses (Table 2i: Appendix I) revealed that the first-year and second-year participants did not differ significantly from one another ( $p = .436$ ). However, the third-year

participants differed significantly from both the first-year participants ( $p = .050$ ) and second-year participants ( $p = .002$ ). Based on the follow-up pairwise comparisons between the IPG and UIPG at each proficiency level (Table 3i: Appendix I), the L2 participant group factor yielded a significant main effect because the UIPG significantly outperformed the IPG at the second-year level ( $p = .016$ ) and near-significantly at the third-year level ( $p = .062$ ). In addition, based on the follow-up pairwise comparisons between all proficiency levels within each L2 group (Table 4i: Appendix I), the L2 proficiency level factor yielded a significant main effect because the third-year IPG participants significantly outperformed only the second-year IPG participants ( $p = .011$ ) and because the third-year UIPG participants significantly outperformed both the first-year UIPG participants ( $p = .019$ ) and second-year UIPG participants ( $p = .025$ ).

To examine whether the control participants differed from the L2 participants representing each level within each L2 group in their accuracy on the verbal present negation, six independent sample  $t$ -tests were performed. The results revealed that only the third-year UIPG participants were statistically comparable with the control participants ( $p = .195$ ), while the other L2 participants of both L2 groups significantly underperformed the controls (Table 4.19).

**Table 4.19** Results of  $t$ -Tests on the Accuracy Mean Scores of the CG and Each Proficiency Level Within Each L2 Group on the Verbal Present negation

Variable (I)	Variable (J)	df	t	Sig. p	95% Confidence Interval	
					Lower Bound	Upper Bound
CG	IPG					
Native (n = 9)	1st year (n = 9)*	8.000	-4.063	.004	-65.40	-18.04
	2nd year (n = 11)*	10.000	-5.216	<.001	-89.81	-36.05
	3rd year (n = 8)*	7.000	-2.538	.039	-56.35	-1.99
CG	UIPG					
Native (n = 9)	1st year (n = 10)*	9.000	-5.819	<.001	-47.66	-20.98
	2nd year (n = 10)*	9.000	-4.996	<.001	-47.74	-17.98
	3rd year (n = 9)*	8.000	-1.414	.195	-9.73	2.33

Note. \* = equality of variances was not assumed via Levene's  $F$  test ( $p < .05$ ).

In sum, the qualitative and quantitative analyses of the verbal present negation data revealed a number of findings. Initially, the IPG exhibited a tendency to use the MSA negation system more often than its EA counterpart. This is evidenced by the fact that only 17.5% of their total tokens exhibited the EA negation system. Interestingly, only two of these EA-like tokens were completely correct and were produced by one first-year participant with the same verb in the 3.s.m context. Additionally, those EA tokens were prevalent in the data of second-year participants, whereas the first-year and third-year IPG participants opted predominantly for the MSA negation system. Furthermore, both L2 groups exhibited a tendency to produce the negation structure (in MSA) at a high accuracy rate almost across all proficiency levels, which suggests that emergence of the MSA negation structure took place from early on. However, the UIPG exhibited a higher accuracy rate, compared to the IPG, evident in the significant and near-significant outperformance of the first-year and third-year UIPG participants over their IPG counterparts. The advantage of the UIPG is further supported by two pieces of evidence. First, the third-year UIPG participants differed significantly from the first-year and second-year UIPG participants, whereas the accuracy mean of the third-year IPG participants was statistically comparable to that of the first-year IPG participants. Second, only the third-year UIPG participants were statistically on par with the control participants. These two pieces of evidence indicate that the IPG participants did not improve significantly in the long run, unlike the UIPG participants, who were able to approximate further toward native-like proficiency over time—at the third-year level.

#### **4.4 Complexity**

This section presents the results related to the complexity construct of all participants representing the IPG, UIPG, and CG. This proficiency dimension was examined particularly in

the acquisition data of the structures of subjunctive subordination and cause–effect subordination, which were elicited in the picture-based manipulation task that was held in the form of in-person one-on-one interviews with all participants (see Section 3.4.2.2: Chapter 3). The results revealed by the data of subjunctive subordination and cause–effect subordination are presented below in Sections 4.4.1 and 4.4.2, respectively.

#### 4.4.1 Subjunctive subordination

Based on the detailed coding criteria discussed in Section 3.6.2.5, the production data of the subjunctive subordination structure were analyzed accordingly. Table 4.20 provides descriptive statistics for the subjunctive subordination data of the IPG, UIPG, and CG. Examples of the L2 participants’ correct tokens of the subjunctive subordination are provided in (99a–f). Recall, in addition to the MSA subjunctive *ʔan* construction, the IPG participants were introduced to the EA equivalent construction, which entails dropping the particle *ʔan* “to” and spelling out the habitual and progressive prefix *bi-* with the matrix verb in the simple and progressive present contexts (i.e., the contexts in which the questions of the subjunctive subordination tasks were asked). Interestingly, out of the 313 subjunctive subordination tokens produced by the IPG, only six tokens were completely correct in the EA variety (0.02%), and they were produced by two first-year participants and one second-year participant. The finite complement/subordinate clauses in these tokens were not subcategorized by the imperfect form of verbs in the matrix clause but rather by the active participles *ʕāyiz/a* “wanting,” *lāzim* “obligating,” and *mumkin* “could be,” as in (100a, b, c), respectively. As for the control participants, they opted more often for subcategorizing the finite subordinate clauses by the imperfect form of verbs preceded by the habitual and progressive prefix *bi-*, as in (101a–b), and less often by the active participles *ʕāyiz/a* “wanting,” *ʔādir/ʔādra* “able,” and *nīfsu/nīfsahā* “wishing,” as in (102a–b).

**Table 4.20** Descriptive Statistics for the Subjunctive Subordination Data of the Participants in the IPG, UIPG, and CG

Participant Group Proficiency Level	Subjunctive Subordination			
	C/T	%	<i>M</i>	<i>SD</i>
<b>IPG</b>				
1st year ( <i>n</i> = 9)	21.5/99	21.7	21.83	12.98
2nd year ( <i>n</i> = 11)	22.5/120	18.8	19.15	23.56
3rd year ( <i>n</i> = 8)	34/94	36.2	35.90	15.55
<b>UIPG</b>				
1st year ( <i>n</i> = 10)	27.5/93	29.6	27.45	21.52
2nd year ( <i>n</i> = 10)	47/111	42.3	41.18	20.44
3rd year ( <i>n</i> = 9)	80/104	76.9	77.27	19.99
<b>CG</b>				
Native ( <i>n</i> = 9)	106/106	100	100	00

Note. C/T = Correct/Total; % of correct tokens; M = Mean; SD = Standard Deviation.

- (99) a. IL: *hiya tu-rīd ʔan ta-ṭār fī ʃaff ʔawwal*  
she Imerf.3.s.f-want to Imerf.3.s.f-fly in class.s.m first.s.m  
TL: *hiya tu-rīd ʔan tu-sāfir fī ʔad-daraj-a ʔal-ʔūlā*  
she Imerf.3.s.f-want to Imerf.3.s.f-travel in the-class-s.f the-first.s.f  
“She wants to travel in first class.”  
(IPG: 1st-year participant)
- b. *huwa yu-ḥibb-u ʔan ya-ṭaxarraj*  
he Imerf.3.s.m-like-Indic to Imerf.3.s.m-graduate  
“He likes to graduate.”  
(UIPG: 1st-year participant)
- c. *ʔar-ragul yu-rīd ʔan ya-ṭazawwaq ʔimraʔa*  
the-man Imerf.3.s.m-want to Imperf.3.s.m-marry woman  
“The man wants to marry a woman.”  
(IPG: 2nd-year participant)
- d. *hiya ta-rfuḍ ʔan tu-daxxin*  
she Imerf.3.s.f-refuse to Imperf.3.s.m-smoke  
“She refuses to smoke.”  
(UIPG: 2nd-year participant)

- e. *yu-ḥāwil*                      *ʔan*    *ya-ʔxuḏ-a*                      *ʔal-ʔimtiḥān*  
 Imerf.3.s.m-try                      to    Imperf.3.s.m-take-Subj                      the-test  
 “He is trying to take the test.”  
 (IPG: 3rd-year participant)
- f. *ʔal-ʔimraʔa*                      *ta-staʔīʔ*                      *ʔan*    *ta-lʔab*                      *mūsīqā*  
 the-woman                      Imerf.3.s.f-can                      to    Imerf.3.s.f-play                      music  
 “The woman can play music.”  
 (UIPG: 3rd-year participant)
- (100) a. IL: *hiya*    *ʔāyiz-a*                      *ti-sāfir*                      *fī*    first                      class  
           she    wanting-s.f    Imerf.3.s.f-travel                      in    first                      class  
 Tl: *hiyya*    *ʔāyz-a*                      *ti-sāfir*                      *fī*    *ʔid-darag-a*                      *ʔil-ʔūlā*  
           she    wanting-s.f    Imerf.3.s.f-travel                      in    the-class.s.f                      the-first.s.f  
 “She wants [to] travel in first class.”  
 (IPG: 1st-year participant)
- b. IL: *huwa*                      *lāzim*                      *yu-sāfir*                      *fī*    *ʔal-maktab*  
           he                      obligating                      Imerf.3.s.m-travel                      in    the-office.s.m  
 Tl: *huwwa*                      *lāzim*                      *yi-rūḥ*                      *li*    *ʔil-maktab-a*  
           he                      obligating                      Imerf.3.s.m-go                      to    the-library-s.f  
 “He [has to go] to the library.”  
 (IPG: 1st-year participant)
- c. IL: *ʔar-rajul*                      *mumkin*                      *yi-rūḥ*                      *fī*    airplane  
           the-man                      could                      Imerf.3.s.m-go                      in    airplane  
 Tl: *ʔir-rāgil*                      *mumkin*                      *yi-sāfir*                      *bi*    *ʔiṭ-ṭayyār-a*  
           the-man                      could                      Imerf.3.s.m-go                      by    the-airplane-s.f  
 “The man could travel by the airplane.”  
 (IPG: 2nd-year participant)
- (101) a. *huwwa*    *ʔammāl*    *bi-yi-ḥāwil*                      *yi-šil*                      *ʔil-wazn*  
           he                      is.s.m    Prog-Imperf.3.s.m-try                      Imperf.3.s.m-lift                      the-weight  
 “He is trying to lift weight.”  
 (control participant)
- b. *ʔis-sitt*                      *bi-ti-rfuḏ*                      *ʔinn*    *hiyya*    *ti-šrab*                      *sagāyir*  
           the-woman    Hb-Imperf.3.s.f-refuse                      that    she    Imperf.3.s.f-drink                      cigarettes.PL  
 “The woman refuses to smoke cigarettes.”  
 (control participant)

- (102) a. *ʔir-rāgil*            *dah*            *ʕāyiz*            *yi-tgawwiz*  
 The-man            this            wanting.s.m            Imperf.3.s.m-marry  
 “This man wants to get married.”  
 (control participant)
- b. *ʔis-sitt*            *nifsa-hā*            *ti-bʔā*            *yaniyy-a*  
 the-woman            wishing-s.f            Imperf.3.s.f-become            rich-s.f  
 “The woman wishes to become rich.”  
 (control participant)

Recall, the subjunctive subordination is considered a complex structure that can be optionally substituted by a verbal noun derived from the subjunctive subordinate verb. Usage of the optional verbal noun construction transforms the subjunctive subordination structure into a simple one. Thus, productions of the optional verbal noun structure were not counted toward the total tokens of the subjunctive subordination. Rather, they were coded separately to examine the extent to which the participants opted for it, in comparison to the complex structure.

The analysis revealed that the verbal noun structure was produced 65 times by all the participants. Table 4.21 provides a snapshot of the correct and incorrect production of this simple structure by all participants. The first-year participants, especially those of the UIPG, opted for this structure the most because they produced the vast majority of it: 9 by the IPG and 27 by the UIPG. The second-year participants produced 21 tokens: 12 by the IPG and 9 by the UIPG. The third-year participants produced the least tokens: 4 by the IPG and 4 by the UIPG. Overall, 52 tokens were completely correct, as in (103a–b), and only 13 tokens were incorrect due to inflecting the produced verbs for the 1.s and 3.s.m forms instead of their 3.s.f counterpart, as in (104 a–b). The distribution of the 13 incorrect tokens among the participants was as follows: 2 by first-year IPG participants, 4 by second-year IPG participants, 1 by a third-year IPG participant, 4 by first-year UIPG participants, and 2 by second-year UIPG participants. This finding indicates that the lower-proficiency L2ers exhibited a tendency to avoid the complex



structure because it perhaps added extra-cognitive load on them, which was not an issue for the advanced L2ers, who produced the complex structure more often than those at the beginner and intermediate proficiency levels. This inference is further supported by the data of the control participants, who produced the simple verbal noun structure only twice, as in (105).

**Table 4.21** Production of the Simple Verbal Noun Structure Instead of the Complex Subjunctive Subordination Structure by the IPG, UIPG, and CG

Participant Group Proficiency Level	Use of the Simple Verbal Noun Structure	
	C/T	%
<b>IPG</b>		
1st year ( <i>n</i> = 9)	7/9	77.8
2nd year ( <i>n</i> = 11)	8/12	66.7
3rd year ( <i>n</i> = 8)	1/2	50
<b>UIPG</b>		
1st year ( <i>n</i> = 10)	23/27	85.2
2nd year ( <i>n</i> = 10)	7/9	77.8
3rd year ( <i>n</i> = 9)	4/4	100
<b>CG</b>		
Native ( <i>n</i> = 9)	2/2	100

Note. C/T = Correct/Total; % of correct tokens.

(103) a. *yu-fakkir*                      *fī*      *ʔas-safar*                      *ʔilā*      *balad*                      *ʔāxar*  
 Imperf.3.s.m-think                      in      the-traveling.VN                      to      country.s.m                      another.s.m  
 “He is thinking about traveling to another country.”

(IPG: 2nd-year participant)

b. *hiya*                      *ta-staʔīf*                      *laʔib-a*                      *ʔal-mūsīqā*  
 she                      Imperf.3.s.f-can                      playing.VN-Acc                      the-music  
 “She can play music.”

(UIPG: 2nd-year participant)

(104) a. IL: *\*hiya*                      *lā*                      *ʔu-rīd-u*                      *ʔat-tadxīn*  
                     she                      does not                      Imperf.1.s-want-Indic                      the-smoking.VN  
                     TL: *hiya*                      *lā*                      *tu-rīd-u*                      *ʔat-tadxīn*  
                     she                      does not                      Imperf.3.s.f-want-Indic                      the-smoking.VN  
 “She does not want [to] smoke.”

(IPG: 1st-year participant)

b. IL: *hiya	ya-rfuḍ	ʔar-raqṣ
she	Imperf.3.s.m-refuese	the-dancing.VN
TL: hiya	ta-rfuḍ	ʔar-raqṣ
she	Imperf.3.s.f-refuese	the-dancing.VN
“She refuses [to] dance.”		

(UIPG: 1st-year participant)

(105) wi	rafīḍ-it	barḍu	ʕarḍ	ʔir-raʔṣ
and	Perf.refused-3.s.f	also	offer	the-dancing.VN
“And she also refused a dancing offer.”				

(control participant)

The qualitative analysis of the participants’ data of the subjunctive subordination structure revealed eight patterns of errors, as summarized in Table 4.22.

**Table 4.22** Error Patterns Produced by the Participants in the Subjunctive Subordination (Error Type Tokens/Total Errors)

Error Type	IPG			UIPG		
	1st year (n = 9)	2nd year (n = 11)	3rd year (n = 8)	1st year (n = 10)	2nd year (n = 10)	3rd year (n = 9)
L of S	31/92	54/103	32/66	38/75	32/69	10/25
Authenticity*	29/92	11/103	12/66	19/75	10/69	2/25
Agr	13/92	11/103	8/66	3/75	4/69	3/25
Authenticity + Agr	13/92	12/103	1/66	7/75	5/69	2/25
Authenticity + T	2/92	8/103	3/66	7/75	2/69	0/25
T	1/92	5/103	3/66	31/75	5/69	4/25
ʔan followed by VN	2/92	0/103	7/66	0/75	7/69	2/25
T + Agr	1/92	2/103	0/66	0/75	4/69	2/25

Note. \* = coded as partially incorrect; L of S = Lack of Suppliance; T = Tense; Agr = Agreement; VN = Verbal Noun.

The most frequent error was due to lack of subjunctive subordination suppliance, which constituted 45.8% of the total errors. It occurred mostly when the participants could not provide descriptions for the target pictures; provided English subordinate verbs instead of the Arabic equivalents, as in (106a); did not provide subordinate verbs at all, as in (106b); and provided simple sentences that contained neither a subjunctive construction nor its optional verbal noun



contexts (see Appendix E), the matrix verb in EA-like productions was expected to be spelled out with the habitual or progressive prefix *-bi*. However, almost all production data of the subjunctive structure by the IPG participants involved matrix and subordinate verbs that are mutually used in both MSA and EA, and the matrix verbs were not preceded by the prefix *-bi*. It was consequently challenging to judge whether those production tokens were in MSA or EA. Therefore, morphosyntactically error-free productions that involved subordinate clauses not headed by the complementizer *ʔan* “to” were considered unauthentic and hence coded as partially incorrect. This coding criterion was adhered to consistently across all IPG and UIPG participants. Similarly, spelling out the complementizer in EA-like productions was considered as an authentic error and hence coded as partially incorrect. Accordingly, the authenticity error constituted 19.3% of the total errors. It occurred when the participants omitted the particle *ʔan* “to,” as in (107a–d), and only one token produced by a first-year IPG participant involved spelling out the particle *ʔan* “to” in a complement clause subcategorized by the active participle *lāzim* “wanting,” which made the whole production in (107e) sound as an EA-like sentence. The lack of *ʔan* “to” suppliance error was prevalent among the first-year participants of both L2 groups. However, it decreased among the second-year IPG participants but maintained a steady status at the third-year level of the IPG. Conversely, this type of error declined steeply as proficiency increased within the UIPG, whose third-year participants produced only two tokens involving omission of the complementizer *ʔan* “to.”

(107)	a.	IL: <i>*huwa yu-rīd</i>		<i>ya-jid</i>		<i>zawj-a</i>
		he Imperf.3.s.m-want		Imperf.3.s.m-find		wife-s.f
		TL: <i>huwa yu-rīd</i>	<i>ʔan</i>	<i>ya-jid</i>		<i>zawj-a</i>
		he Imperf.3.s.m-want	to	Imperf.3.s.m-find		wife-s.f
		“He wants to find a wife.”				

(IPG: 2nd-year participant)

b. IL: \**huwa yu-fakkir* *ya-drus* *fī* *ʔal-maktab-a*  
 he Imperf.3.s.m-think Imperf.3.s.m-study in the-library-s.f  
 TL: *huwa yu-fakkir* *ʔan ya-drus* *fī* *ʔal-maktab-a*  
 he Imperf.3.s.m-think Imperf.3.s.m-study in the-library-s.f  
 “He is thinking to study at the library.”

(UIPG: 1st-year participant)

c. IL: \**wa ta-ryab* *ta-ʕmal* *fī* *ʔal-kumbiyūtar*  
 and Imperf.3.s.f-want Imperf.3.s.m-work in the-computer  
 TL: *wa ta-ryab* *ʔan ta-ʕmal* *ʕalā* *ʔal-kumbiyūtar*  
 and Imperf.3.s.f-want to Imperf.3.s.m-work on the-computer  
 “And she wants to work on the computer.”

(IPG: 3rd-year participant)

d. IL: \**huwa yu-ḥāwil* *ya-ktub* *ʔal-ʔimtiḥān*  
 he Imperf.3.s.m-try Imperf.3.s.m-write the-test  
 TL: *huwa yu-ḥāwil* *ʔan ya-ktub* *ʔal-ʔimtiḥān*  
 he Imperf.3.s.m-try to Imperf.3.s.m-write the-test  
 “He is trying to write the test.”

(UIPG: 3rd-year participant)

e. IL: \**huwa lāzim* *ʔan ya-xarraǰ* *ʔal-ʔimtiḥān*  
 he must to Imperf.3.s.m-[finish] the-test  
 TL: *huwwa lāzim* *yi-xallaṣ* *ʔal-ʔimtiḥān*  
 he must Imperf.3.s.m-finish the-test  
 “He has [to] finish the test.”

(IPG: 1st-year participant)

The third frequent error was due to inflecting the subordinate verbs mostly or matrix verbs rarely for the wrong agreement. It constituted 9.8% of the total errors and occurred when verbs were inflected for the 3.s.m form instead of its 3.s.f counterpart, as in (108a–b), and when they were inflected for the 1.s form instead of its 3.s.m or 3.s.f counterparts, as in (108c–d), respectively. The IPG participants exhibited a tendency to produce this error more often than their UIPG counterparts across all proficiency levels. It was in fact more prevalent in the data of the third-year IPG participants, compared to the data of the first-year UIPG participants. As for the wrong



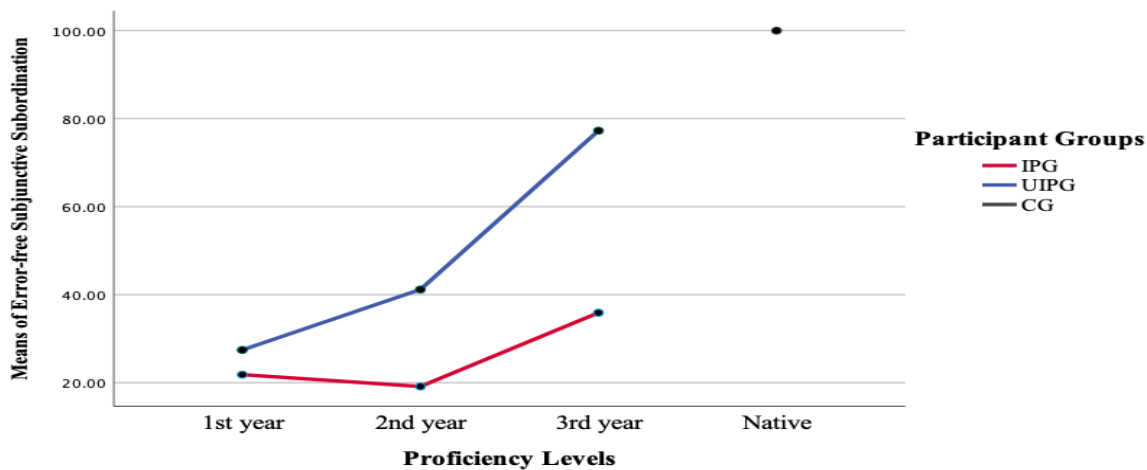


	b. IL: * <i>hiya ta-staṭiṣ-u</i>	<i>ʔan</i>	<i>jarī</i>
	she Imperf.3.s.f-can-Indic	to	running.VN
	TL: <i>hiya ta-staṭiṣ-u</i>	<i>ʔan</i>	<i>ta-jrī</i>
	she Imperf.3.s.f-can-Indic	to	Imperf.3.s.f-run
	“She can run.”		
			(UIPG: 2nd-year participant)
(112)	IL: * <i>hiya muš tu-rīd</i>	<i>ʔan</i>	<i>daxxan</i>
	he Neg Imperf.3.s.m-want	to	Perf.smoked.3.s.m
	TL: <i>hiya lā tu-rīd</i>	<i>ʔan</i>	<i>tu-daxxin</i>
	he Neg Imperf.3.s.m-want	to	Imperf.3.s.f-smoke
	“She does not want to smoke.”		
			(IPG: 2nd-year participant)

The aggregate error-free production data of the subjunctive subordination reveal an advantage for the UIPG over the IPG (see Table 4.20 above and Figure 4.9 below). Overall, the mean scores of the IPG were lower than those of the UIPG across all proficiency levels. Interestingly, the mean score of the second-year UIPG participants was even higher than that of the third-year IPG participants. These observations may indicate that significant differences exist between the two L2 groups. Moreover, as depicted in Figure 4.9, the mean scores of the IPG and UIPG were not parallel. The UIPG performance (the blue line) increased quite steeply from the first-year to third-year levels, whereas the IPG performance (the red line) did not exhibit such a considerable improvement. This observation points to an interaction between the two independent variables of L2 participant group and proficiency level. The development from the first-year to second-year to third-year levels in the error-free production of the subjunctive subordination was contingent upon the type of program to which the participants belonged. In other words, only those of the UIPG improved considerably in their acquisition of this structure as their proficiency increased.



**Figure 4.9** Means of the IPG, UIPG, and CG on the Error-Free Production of the Subjunctive Subordination



The above-mentioned observations were confirmed by the full factorial two-way ANOVA test that was conducted to examine the main effects of the two independent variables (the L2 participant group and L2 proficiency level), as well as their interaction effects, on the error-free production of the subjunctive subordination by the two L2 groups across the three proficiency levels. The main effect of the L2 participant group factor was statistically significant ( $F(1,51) = 19.312, p < .001$ ) and yielded an effect size of .275, indicating that 27.5% of the variances in the L2 participants' means were explained by the L2 participant groups to which they belonged. The main effect of the L2 proficiency level factor was likewise significant ( $F(2,51) = 13.367, p < .001$ ) and yielded an effect size of .344, indicating that 20% of the variances in the L2 participants' means were explained by their proficiency levels. The interaction effect was also significant ( $F(2,51) = 3.698, p = .032, \text{partial } \eta^2 = .127$ ), indicating that the main effect of the L2 participant group factor depends on the L2 proficiency level factor and vice versa. Tukey's post hoc analyses (Table 1j: Appendix J) showed that the first-year and

second-year participants did not differ significantly from each other ( $p = .718$ ). However, the third-year participants differed significantly from both the first-year participants ( $p < .001$ ) and second-year participants ( $p < .001$ ). Based on the follow-up pairwise comparisons between the IPG and UIPG at each proficiency level (Table 2j: Appendix J), the L2 participant group factor yielded a significant main effect due to the significant outperformance of the UIPG over the IPG at the second-year level ( $p = .013$ ) and third-year level ( $p < .001$ ). In addition, based on the follow-up pairwise comparisons between all the proficiency levels within each L2 group (Table 3j: Appendix J), the main effect of the L2 proficiency level factor was significant because the third-year UIPG participants significantly outperformed both the first-year UIPG participants ( $p < .001$ ) and second-year UIPG participants ( $p < .001$ ). In contrast, the IPG participants were statistically comparable across all proficiency levels.

To examine whether the CG differed from the L2 participants representing each proficiency level within each L2 group in their error-free production of the subjunctive subordination, six independent sample  $t$ -tests were performed. The results revealed that the CG significantly outperformed the IPG and UIPG across the three proficiency levels (Table 4.23).

**Table 4.23** Results of  $t$ -Tests on the Means of the CG and Each Proficiency Level Within Each L2 Group on the Error-Free Production of Subjunctive Subordination

Variable (I)	Variable (J)	$df$	$t$	Sig. $p$	95% Confidence Interval	
					Lower Bound	Upper Bound
CG	IPG					
Native ( $n = 9$ )	1st year ( $n = 9$ )*	8.000	-18.067	< .001	-88.14	-68.19
	2nd year ( $n = 11$ )*	10.000	-11.384	< .001	-96.68	-65.03
	3rd year ( $n = 8$ )*	7.000	-11.662	< .001	-77.10	-51.10
CG	UIPG					
Native ( $n = 9$ )	1st year ( $n = 10$ )*	9.000	-10.661	< .001	-87.94	-57.16
	2nd year ( $n = 10$ )*	9.000	-9.101	< .001	-73.44	-44.20
	3rd year ( $n = 9$ )*	8.000	-3.411	.005	-38.10	-7.37

Note. \* = equality of variances was not assumed via Levene's  $F$  test ( $p < .05$ ).

In sum, the qualitative analysis of the subjunctive subordination data revealed several findings. First, the IPG exhibited a tendency once again to produce the MSA construction more often than its EA counterpart. Indeed, only six tokens produced by two first-year participants and one second-year participants were completely correct in the EA variety, and the subordinate clauses in those six tokens were not subcategorized by imperfect verbs in the matrix clause but rather by the active participles *řāyız/a* “wanting,” *lāzim* “obligating,” and *mumkin* “could be.” Second, the error pertaining to lack of *řan* suppliance decreased significantly over time, particularly at the third-year level of the UIPG. It also showed a decreasing behavior from the first-year to second-year levels of the IPG only because the number of this error produced by the third-year IPG participants was equal to that produced by the second-year IPG participants. Third, most of the produced errors were in the subordinate clause, indicating that the L2 participants started the sentence properly; yet and perhaps due to the cognitive load, they made various errors when they attempted to embed the subordinate clause within the matrix clause. As for the errors made in the matrix clauses, they were mostly produced by the IPG participants, who used the active participle *řāyız* “wanting.m” in the 3.s.f context.

As for the quantitative analysis of the error-free subjunctive subordination data, it revealed that the UIPG significantly outperformed the IPG at the second-year and third-year levels. The advantage of the former was further evidenced by the fact the third-year UIPG participants significantly outperformed the first-year and second-year UIPG participants, whereas the IPG participants were statistically comparable across all proficiency levels. Nevertheless, the significant outperformance of the UIPG was not sufficient when compared to the CG, whose error-free production of the subjunctive subordination was at a significantly higher rate than both L2 groups across all proficiency levels.

#### 4.4.2 Cause–effect subordination

Based on the detailed coding criteria discussed in Section 3.6.2.6, the acquisition data of the cause–effect subordination structure were analyzed accordingly. Table 4.24 provides descriptive statistics for the cause–effect subordination data of the IPG, UIPG, and CG. Examples of the L2 participants’ correctly produced tokens are provided in (113a–d). Recall, in addition to the conjunction *liʔann* “because,” the IPG participants were introduced to the EA conjunction *ʕašān/ʕalašān* “because,” which behaves differently from the former (see Section 3.3.4: Chapter 3). Interestingly, out of all the 214 tokens produced by the IPG participants, only 11 tokens were completely correct when using the conjunction *ʕašān/ʕalašān* “because”: five by first-year participants, three by second-year participants, and three by third-year participants. More remarkably, the subordinate clauses of these correct tokens were subcategorized by adjectives in the matrix clauses, as in (114a–b). As for the control participants, they exhibited a tendency to convey the cause–effect relation more often by the conjunction *ʕašān/ʕalašān* (in 60 tokens) and less often by the conjunction *liʔann* (in 12 tokens), as shown in (115a–b), respectively.

**Table 4.24** Descriptive Statistics for the Cause-Effect Subordination Data of the IPG, UIPG, and CG

Participant Group Proficiency Level	Cause-effect Subordination			
	C/T	%	<i>M</i>	<i>SD</i>
<b>IPG</b>				
1st year ( <i>n</i> = 9)	12.5/72	17.4	17.37	17.62
2nd year ( <i>n</i> = 11)	9.5/81	11.7	12.50	21.65
3rd year ( <i>n</i> = 8)	9.5/61	15.6	15.30	19.20
<b>UIPG</b>				
1st year ( <i>n</i> = 10)	22/65	33.8	33.46	31.86
2nd year ( <i>n</i> = 10)	27/72	37.5	35.71	27.50
3rd year ( <i>n</i> = 9)	46/66	69.7	70.16	15.96
<b>CG</b>				
Native ( <i>n</i> = 9)	72/72	100	100	00

*Note.* C/T = Correct /Total; % of correct tokens; *M* = Mean; *SD* = Standard Deviation.

- (113) a. *ʔar-ragul taʕbān liʔanna ʔal-jaww ḥārr*  
 the-man tired.s.m because the-weather.s.m hot.m  
 “The man is tired because the weather is hot.”  
 (IPG: 2nd-year participant)
- b. *ʔal-bint saʕīd-a liʔanna-hā ʔakal-at ice-cream*  
 the-girl happy-s.f because-she Perf.ate-3.s.f ice-cream  
 “The girl is happy because she ate ice-cream.”  
 (UIPG: 2nd-year participant)
- c. *ʔal-ʔimrʔa ḏahab-at ʔilā ʔad-ductūr liʔanna-hā marīd-a*  
 the-woman Perf.went-3.s.f to the-doctor because-she sick-s.f  
 “The woman went to the doctor because she is sick.”  
 (IPG: 3rd-year participant)
- d. *ya-šʕur ʔal-walad bi ʔal-ḥuzn liʔanna-hu*  
 Imperf.3.s.m-feel the-boy by the-sadness because-he  
*faqada luʕb-at-a-hu*  
 Perf.lost.3.s.m toy-s.f-Acc-his  
 “The boy feels sad because he lost his toy.”  
 (UIPG: 3rd-year participant)
- (114) a. *hiya bārid-a giddan ʕašān ʔaṭ-ṭaqs bārid giddan*  
 she cold-s.f very because the-weather.s.m cold.m very  
 “She is very cold because the weather is very cold.”  
 (IPG: 1st-year participant)
- b. *ʕašān ʔal-gaww kān maṭar hiya zaʕlān-a ʔalʔān*  
 because the-weather.s.m Perf.was.3.s.m rain she sad-s.f now  
 “because the weather was rainy, she is sad now.”  
 (IPG: 2nd-year participant)
- (115) a. *huwwa bi-yi-ʕayyaṭ ʕašān ʔil-liʕb-a bitaṭ-t-u ʔitʔaṭaṭ-it*  
 he Prog-Imperf.3.s.m-cry because the-toy-s.f own-f-his Perf.chopped-3.s.f  
 “He is crying because his toy [got] chopped.”  
 (control participant)
- b. *ʔir-rāgil rigl-u maksūr-a liʔann huwwa wiʔiṭ min ʔis-sillim*  
 the-man foot.f-his broken-s.f because he Perf.fell.3.s.m from the-ladder  
 “The man [has] a broken leg because he fell of the ladder.”  
 (control participant)

Recall, the cause–effect subordination headed by the conjunctions *liʔann* or *ʕašān/ʕalašān* “because” can be optionally substituted by an *ʔidāfah* phrase, headed by the phrase *bisabab* or the EA conjunction *ʕašān* and followed by a genitive noun to convey the meaning of *because of*. Using the optional *ʔidāfah* construction transforms the complex cause–effect construction into a simple one. Thus, production tokens of the simple *ʔidāfah* construction were not counted toward the total tokens of the cause–effect subordination structure. Rather, they were coded separately to examine the extent to which the L2 participants opted for it, as opposed to the complex structure.

The analysis revealed that the L2 participants produced the *ʔidāfah* construction 39 times. Interestingly, the IPG participants opted almost always for the MSA *ʔidāfah* construction headed by *bisabab* “because of” and produced the EA counterpart headed by *ʕašān* “because of” only once. Table 4.25 provides a snapshot of the correct and incorrect production of the optional simple *ʔidāfah* construction by all the participants. The first-year UIPG participants opted for this construction the most since they produced 15 tokens of it. The second-year participants of both L2 groups also produced 15 tokens: seven by the IPG and eight by the UIPG. The third-year participants produced nine tokens only: three by the IPG and six by the UIPG. Moreover, 28 tokens were fully correct, as in (116a–b), and only 11 tokens were incorrect due to the wrong agreement in the 3.s.f context, as in (117 a–b). The distribution of the 11 incorrect tokens of the MSA *ʔidāfah* construction among all the participants was as follows: five by first-year UIPG participants, one by a second-year IPG participant, two by second-year UIPG participants, two by third-year IPG participants, and one by a third-year UIPG participant. The indication of this finding is that advanced L2ers exhibited a tendency to produce the complex structure more often than those at the beginner and intermediate proficiency levels, who opted for the simple structure

in some instances because the complex structure added perhaps extra-cognitive load on them. This assumption is supported by the data of the control participants, who did not produce the optional *ʔiḍāfah* construction at all, and who conveyed the cause–effect relation using the complex structure in all of their description data of the target pictures.

**Table 4.25** Production of the Simple *ʔiḍāfah* construction Instead of the Cause–Effect Subordination Structure by the IPG, UIPG, and CG

Participant Group Proficiency Level	Use of the simple <i>ʔiḍāfah</i> construction	
	C/T	%
<b>IPG</b>		
1st year ( <i>n</i> = 9)	0/0	–
2nd year ( <i>n</i> = 11)	6/7	85.7
3rd year ( <i>n</i> = 8)	1/3	33.3
<b>UIPG</b>		
1st year ( <i>n</i> = 10)	10/15	66.7
2nd year ( <i>n</i> = 10)	6/8	75
3rd year ( <i>n</i> = 9)	5/6	83.3
<b>CG</b>		
Native ( <i>n</i> = 9)	0/0	–

Note. C/T = Correct/Total; % of correct tokens.

- (116) a. *huwa fī kursī mutaḥarrrik bisabab ḥādiθ-a kabīr-a*  
 he in chair.s.m moving.s.m because of accident-s.f big-s.f  
 “He is in a wheel-chair because of a big accident.”  
 (UIPG: 3rd-year participant)
- b. *hiya mubtall-a bisabab ʔal-maṭar*  
 she wet-s.f because of the-rain  
 “She is wet because of the rain.”  
 (IPG: 3rd-year participant)
- (117) a. IL: *\*hiya laysa saḥīd-a bisabab ʔal-jaww*  
 she is not.3.s.m happy-s.f because of the-rain.s.m  
 TL: *hiya lays-at saḥīd-a bisabab ʔal-jaww*  
 she is not-3.s.f happy-s.f because of the-rain.s.m  
 “She is not happy because of the weather.”  
 (UIPG: 1st-year participant)

b. IL: *hiya	ya-ḥīš	bārid	ḥašān	ʔaθ-θalj
she	Imperf.3.s.m-live	cold.s.m	because of	the-snow
TL: hiyya	bardān-a	ḥašān	ʔit-talg	
she	cold-s.f	because of	the-snow	

“She is cold because of the snow.”

(IPG: 2nd-year participant)

The cause–effect subordination data exhibited nine patterns of errors, which are summarized in Table 4.26. The most frequent error was due to lack of cause–effect suppliance, constituting 24.5% of the total errors. It occurred mostly when the participants could not provide descriptions for the target pictures; provided English conjunctions and verbs instead of their Arabic equivalents, as in (118a–b); and produced fragment subordinate clauses, as in (118c–d). The productions here did not convey fully comprehended meanings in Arabic, so they were coded as lack of suppliance for the cause–effect structure.

**Table 4.26** Error Patterns Produced by the Participants in the Cause–Effect Subordination (Error Type Tokens/Total Errors)

Error Type	IPG			UIPG		
	1st year (n = 9)	2nd year (n = 11)	3rd year (n = 8)	1st year (n = 10)	2nd year (n = 10)	3rd year (n = 9)
L of S	15/63	16/72	10/52	10/43	19/45	4/20
VS/NS inst of <i>ʔiḍāfah</i>	3/63	19/72	8/52	8/43	4/45	5/20
VS/NS inst <i>ʔiḍāfah</i> + Agr	3/63	23/72	6/52	6/43	4/45	0/20
Agr in SC	8/63	4/72	2/52	13/43	7/45	7/20
Agr in MC	11/63	1/72	5/52	2/43	4/45	2/20
VS after <i>liʔann</i>	7/63	0/72	9/52	1/43	6/45	1/20
Authenticity + Agr	4/63	2/72	7/52	0/43	0/45	0/20
Authenticity*	7/63	3/72	1/52	0/43	0/45	0/20
Agr in both MC and SC	3/63	0/72	2/52	3/43	0/45	1/20
Neg in either MC or SC	2/63	4/72	2/52	0/43	1/45	0/20
VS after <i>liʔann</i> + Agr	1/63	0/72	2/54	0/43	1/45	0/20

*Note.* \* = coded partially incorrect; L of S = Lack of Suppliance; Agr = Agreement; VS = Verbal Sentence; NS = Nominal-verbless Sentence or Nominal Sentence containing verb predicate; MC = Matrix Clause; SC = Subordinate Clause; Neg = Negation.





the wrong agreement, as in (120a–b). These two patterns of errors constituted, respectively, 15.8% and 14.1% of the total errors produced.

- (119) a. IL: \**huwa fī ʔal-ḥādiθ bisabab ḏahaba bisurʕ-a*  
 he in the-accident.s.m because of Perf.went.3.s.m fast-s.f  
 TL: *huwa fī ʔal-ḥādiθ liʔanna-hu qāda bisurʕ-a*  
 he in the-accident.s.m because-he Perf.drove.3.s.m fast-s.f  
 “He [got] into the accident because he drove fast.”

(UIPG: 1st-year participant)

- b. IL: \**hiya fī ʔal-mayy-a bisabab hiya fī maṭar*  
 she in the-water-s.f because of she in rain  
 TL: *hiyya yarʔān-a mayy-a ʕašān hiyya kān-it fī ʔil-maṭar*  
 she wet-s.f water-s.f because she Perf.was-3.s.f in the-rain  
 “She is wet because she was in the rain.”

(IPG: 2nd-year participant)

- (120) a. IL: \**ʔal-bint saʕīd bisabab ta-ʔkul ice-cream*  
 the-girl happy.s.m because of Imperf.3.s.f-eat ice-cream  
 TL: *ʔal-bint saʕīd-a liʔanna-hā ta-ʔkul ice-cream*  
 the-girl happy-s.f because-she Imperf.3.s.f-eat ice-cream  
 “The girl is happy because she is eating ice-cream.”

(UIPG: 2nd-year participant)

- b. IL: \**ʔal-ibn ya-bkī bisabab ʔal-luʕb-a laysa jayyid*  
 the-son Imperf.3.s.m-cry because of the-toy-s.f is not.3.s.m good.s.m  
 TL: *ʔal-ibn ya-bkī liʔanna ʔal-luʕb-a lays-at jayyidd-a*  
 the-son Imperf.3.s.m-cry because the-toy-s.f is not-3.s.f good.s.f  
 “The son is crying because the toy is not good.”

(IPG: 3rd-year participant)

The fourth most frequent pattern of errors was due to agreement. It occurred mostly when the 3.s.m form was produced instead of the of 3.s.f form in the subordinate clause (13.8% of the total errors), in the matrix clause (8.5%), or in both clauses (1.4%), as in (121a, b, c), respectively.

Another frequent error was due to generating verbal sentences (i.e., VSO word order) instead of nominal sentences (i.e., SVO or nominal-verbless sentence), following *liʔanna* “because,” as in

(122). This error constituted 8.1% of the total errors, and, in some instances, it was accompanied by an agreement error as well (1.4% of the total errors), as in (123).

- (121) a. IL: \**ʔal-marʔa* *ðahab-at* *ʔilā* *ʔad-duktūr* *liʔanna-hā*  
the-woman Perf.went-3.s.f to the-doctor.s.m because-she  
*šaʕar-at* *marīḏ*  
Perf.felt-3.s.f sick.s.m  
TL: *ʔal-marʔa* *ðahab-at* *ʔilā* *ʔad-duktūr* *liʔanna-hā*  
the-woman Perf.went-3.s.f to the-doctor.s.m because-she  
*kān-at* *marīḏ-a*  
Perf.was-3.s.f sick-s.f  
“The woman went to the doctor because she was sick.”

(UIPG: 3rd-year participant)

- b. IL: \**hiya miš mabsūt* *ʕašān* *hiya* *kān-at* *fī* *ʔal-maṭar*  
she is not happy.s.m because she Perf.was-3.s.f in the-rain  
TL: *hiyya miš mabsūt-a* *ʕašān* *hiyya* *kān-it* *fī* *ʔil-maṭar*  
she is not happy-s.f because she Perf.was-3.s.f in the-rain  
“*She is not happy because she was in the rain.*”

(IPG: 2nd-year participant)

- c. IL: \**hiya* *kuwayyis* *giddan* *liʔanna* *wālid-a-hu*  
she good.s.m very because father.m-Acc-his  
*ʕinda-hā* *ḥalawiyy-āt*  
with-her candies-Pl.f  
TL: *hiyya* *mabsūt-a* *giddan* *liʔann* *mām-it-hā*  
she happy-s.f very because mother-s.f-her  
*ʕinda-hā* *ḥalawiyy-āt*  
with-her candy-Pl.f  
“*She is very happy because her mother [has] candies.*”

(IPG: 1st-year participant)

- (122) IL: \**huwa* *zaʕlān* *liʔanna* *qaṭaʕa* *ʔad-dubb*  
he sad.s.m because Perf.chopped.3.s.f the-teddy bear  
TL: *huwwa* *zaʕlān* *liʔann-u* *ʔaṭaʕ* *ʔid-dubb*  
he sad.s.m because-he Perf.chopped.3.s.f the-teddy bear  
“*He is sad because he chopped the teddy bear.*”

(IPG: 3rd-year participant)

(123) IL: \**hiya saḥīd liḡann ḥaṣal-t ḡat-ṭaṣām*  
 she happy.s.m because Perf.got-1.s the-food  
 TL: *hiya saḥīd-a liḡann-hā ḥaṣal-at ḡalā ḡat-ṭaṣām*  
 she happy-s.f because-she Perf.got-3.s.f on the-food  
 “She happy because she got food.”

(IPG: 2nd-year participant)

Another pattern of errors was due to authenticity. As discussed in Section 3.6.2.6 in Chapter 3, any production of *ḡaṣān/ḡalaṣān* in MSA-like sentences that exhibits MSA exclusive verbs (e.g., *yuṣāhid* “to watch,” *yaḏhab* “to go”) renders the sentence unauthentic and hence incorrect, as in (124a). However, if *ḡaṣān/ḡalaṣān* was in utterances that have mutually used verbs in MSA and EA, then the matrix verbs had to be attached to the prefix *-bi* in the habitual present tense and progressive present tense contexts. Otherwise, the production tokens were considered unauthentic, as in (124b). Accordingly, the authenticity errors were coded as partially incorrect and constituted 3.7% of the total errors produced. Nevertheless, the authenticity error was sometimes accompanied by an agreement error, constituting 4.4% of the total errors, as in (125a–b). Thus, the latter tokens were coded as completely incorrect.

(124) a. IL: \**huwa ḡinda-hu ḥādiḡ sayyar-a ḡaṣān huwa ya-ḏhab bisurḡ-a*  
 he with-him accident.s.m car because he Imperf.3.s.m-go fast-s.f  
 TL: *huwwa kān fī ḥādsi-t ḡarabiyy-a ḡaṣān huwwa*  
 he Perf.was.3.s.m in accident-s.f car-s.f because he  
*kān sāyiḡ bisurḡ-a*  
 Perf.was.3.s.m driving.s.m fast-s.f  
 “He was in a car accident because he was driving fast.”

(IPG: 1st-year participant)

b. IL: \**ḡar-rajul ya-ftaḡ ḡaṣ-ṣubbāk ḡaṣān ḡal-jaww ḥārr*  
 the-man Imperf.3.s.m-open the-window because the-weather.s.m hot.s.m  
 TL: *ḡir-rāḡil bi-yi-ftaḡ ḡaṣ-ṣubbāk ḡaṣān ḡal-gaww ḥārr*  
 the-man Prog-Imperf.3.s.m-open the-window because the-weather.s.m hot.s.m  
 “The man is opening the window because the weather is hot.”

(IPG: 3rd-year participant)

- (125) a. IL: \**hiya ta-ḏhab ʔilā ʔad-duktūr-a ʕašān hiya marīḏ*  
 she Imperf.3.s.f-go to the-doctor-s.f because she sick.s.m  
 TL: *hiyya rāḥ-it li ʔad-duktūr-a ʕašān hiyya ʕayyān-a*  
 she Perf.went-3.s.f to the-doctor-s.f because she sick-s.f  
 “She went to the female doctor because she was sick.”  
 (IPG: 1st-year participant)

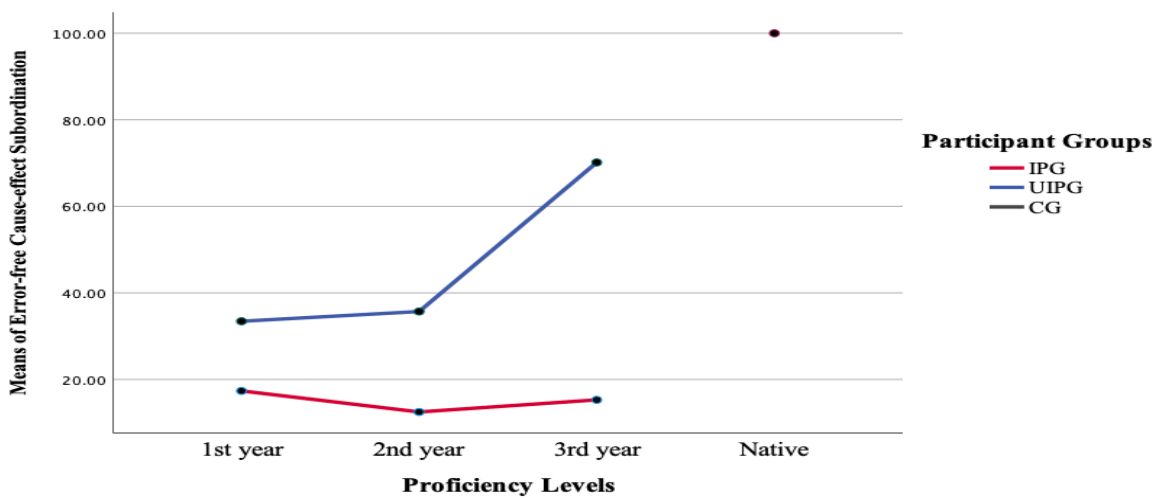
- b. IL: \**sayyār-a ʔar-rajul ya-ṣdum ʔaš-šajar-a ʕašān huwa*  
 car-s.f the-man Imperf.3.s.m-crash the-tree-s.f because he  
*lā ya-qraʔ ʔal sign*  
 does not Imperf.3.s.m-read the sign  
 TL: *sayyār-at ʔir-rāgil xabaṭ-it ʔiṣ-šagar-a ʕašān huwwa*  
 car-s.f the-man Perf.hit-3.s.f the-tree-s.f because he  
*ma-ʔarā-š ʕalām-it ʔis-surʕ-a*  
 Neg-Perf.read.3.s.m-Neg sign-s.f the-speed-s.f  
 “The man’s car hit the tree because he did not read the speed sign.”  
 (IPG: 3rd-year participant)

The last type of error was due to producing a wrong negation construction in either the matrix clause or subordinate clause, as in (126). This error constituted 3.1% of the total errors and was mostly produced by the IPG participants. This finding is similar to the previous findings revealed by the data of past and present negation. That is, the prevalence of this error in IPG participants’ tokens in the cause–effect subordination structure adds support to the finding that the IPG participants significantly underperformed their UIPG counterparts in their accuracy on the negation structures.

- (126) IL: \**huwa miš kuwayyis liʔanna-hu huwa lā*  
 he is not good.s.m because-he he does not  
*laʕaba maʕa toy*  
 Perf.played.3.s.m with toy  
 TL: *huwwa miš mabsūṭ liʔann-u ma-lʕib-š*  
 he is not happy.s.m because-he Neg-Perf.played.3.s.m-Neg  
*bi liʕbi-t-u*  
 by toy-s.f-his  
 “He is not happy because he did not play with his toy.”  
 (IPG: 2nd-year participant)

The aggregate error-free production data of the cause–effect subordination shows again an advantage for the UIPG over the IPG (see Table 4.24 above and Figure 4.10 below). Overall, the means of the IPG participants were lower than those of their UIPG counterparts across all proficiency levels. In fact, the means of the first-year and second-year UIPG participants were even higher than that of the third-year IPG participants. These observations point to significant differences between the two L2 groups. Moreover, as illustrated by Figure 4.10, the scores of the IPG and UIPG were not parallel as well. The performance of UIPG (the blue line) increased steeply from the first-year and second-year levels to the third-year level, whereas the IPG (the red line) was flat across all proficiency levels. This observation indicates an interaction between the two independent variables of the L2 participant group and L2 proficiency level, as improvement in the error-free production of the cause–effect subordination depends on the type of program to which the participants belonged. In other words, only those of the UIPG improved considerably in their acquisition of this complex structure as their proficiency increased.

**Figure 4.10** Means of the IPG, UIPG, and CG on the Error-Free Production of the Cause–effect Subordination



The above-mentioned visual observations were confirmed by the full factorial two-way ANOVA test that was carried out to examine the main effects of the two independent variables (the L2 participant group and L2 proficiency level), as well as their interaction effects, on the error-free production of the cause-effect subordination by the two L2 groups across the three proficiency levels. The main effect of the L2 participant group factor was statistically significant ( $F(1,51) = 25.553, p < .001$ ) and yielded an effect size of .334, indicating that 33.4% of the variances in the L2 participants' means were explained by the participant groups to which they belonged. The main effect of the L2 proficiency level factor was likewise significant ( $F(2,51) = 3.559, p = .036$ ) and yielded an effect size of .122, indicating that 12.2% of the variances in the L2 participants' means were explained by their proficiency levels. The interaction effect was also significant ( $F(2,51) = 3.470, p = .039, \text{partial } \eta^2 = .120$ ), indicating that the main effect of the L2 participant group factor depends on the L2 proficiency level factor and vice versa. Tukey's post hoc analyses (Table 4j: Appendix J) did not reveal any significant differences between the first-year and second-year participants ( $p = .949$ ). However, the third-year participants differed near-significantly from the first-year participants ( $p = .054$ ) and significantly from the second-year participants ( $p = .023$ ). Based on the follow-up pairwise comparisons between the IPG and UIPG at each proficiency level (Table 5j: Appendix J), the L2 participant group factor was significant due to the outperformance of the UIPG over the IPG at the second-year level ( $p = .027$ ) and third-year level ( $p < .001$ ). Additionally, based on the follow-up pairwise comparisons between all proficiency levels within each L2 group (Table 6j: Appendix J), the L2 proficiency level factor was significant because the third-year UIPG participants significantly outperformed both the first-year UIPG participants ( $p = .001$ ) and second-year UIPG participants ( $p = .002$ ),

whereas the mean scores of the IPG participants were statistically comparable across the three proficiency levels.

To examine whether the control participants differed from the L2 participants of each proficiency level within each L2 group in their error-free production of the cause–effect subordination, six independent sample *t*-tests were performed. The results revealed that the CG significantly outperformed both the IPG and UIPG across all proficiency levels (Table 4.27).

**Table 4.27** Results of *t*-Tests on the Mean Scores of the CG and Each Proficiency Level Within Each L2 Group on the Error-Free Production of the Cause–Effect Subordination

Variable (I)	Variable (J)	<i>df</i>	<i>t</i>	Sig. <i>p</i>	95% Confidence Interval	
					Lower Bound	Upper Bound
CG	IPG					
Native ( <i>n</i> = 9)	1st year ( <i>n</i> = 9)*	8.000	-14.072	<.001	-96.17	-69.09
	2nd year ( <i>n</i> = 11)*	10.000	-13.404	<.001	-102.05	-72.95
	3rd year ( <i>n</i> = 8)*	7.000	-12.476	<.001	-100.75	-68.65
CG	UIPG					
Native ( <i>n</i> = 9)	1st year ( <i>n</i> = 10)*	9.000	-6.605	<.001	-89.33	-43.75
	2nd year ( <i>n</i> = 10)*	9.000	-7.392	<.001	-83.96	-44.62
	3rd year ( <i>n</i> = 9)*	8.000	-5.611	<.001	-42.11	-17.58

Note. \* = equality of variances was not assumed via Levene’s *F* test (< .05).

Taken altogether, the qualitative and quantitative analyses of the cause–effect subordination data revealed several observations. Initially, the IPG participants exhibited a tendency to use the conjunctions *li?anna* and *fašān* “because” interchangeably. However, they produced the former more accurately than the latter. This is evidenced by their unauthentic use of the conjunction *fašān* in MSA-like sentences. In fact, only six tokens produced by two first-year participants and one second-year participant were completely correct in the EA variety, and the complement clauses in these tokens were not subcategorized by imperfect verbs in the matrix clause but rather by the active participles *fa?yiz/a* “wanting,” *lāzim* “obligating,” and *mumkin*



“could be.” In addition, one of the most prevalent error patterns pertained to producing verbal, nominal-verbal, or nominal-verbless sentences in subordinate clauses headed by *bisabab* “because of,” which is the first part of the *ʔidāfah* construction. This error was prevalent in the data, particularly those of the IPG. Moreover, although the first-year participants of both L2 groups were statistically comparable, the second-year and third-year UIPG participants significantly outperformed their IPG counterparts. The advantage of the UIPG over the IPG also stems from the within-group analysis, revealing that the IPG did not exhibit any statistically significant improvement from the first-year to second-year to third-year levels, whereas the UIPG participants improved significantly in the long run, particularly at the third-year level. However, despite the significant outperformance of the UIPG over the IPG, the mean score of the CG was significantly higher than those of both L2 groups across the three proficiency levels.

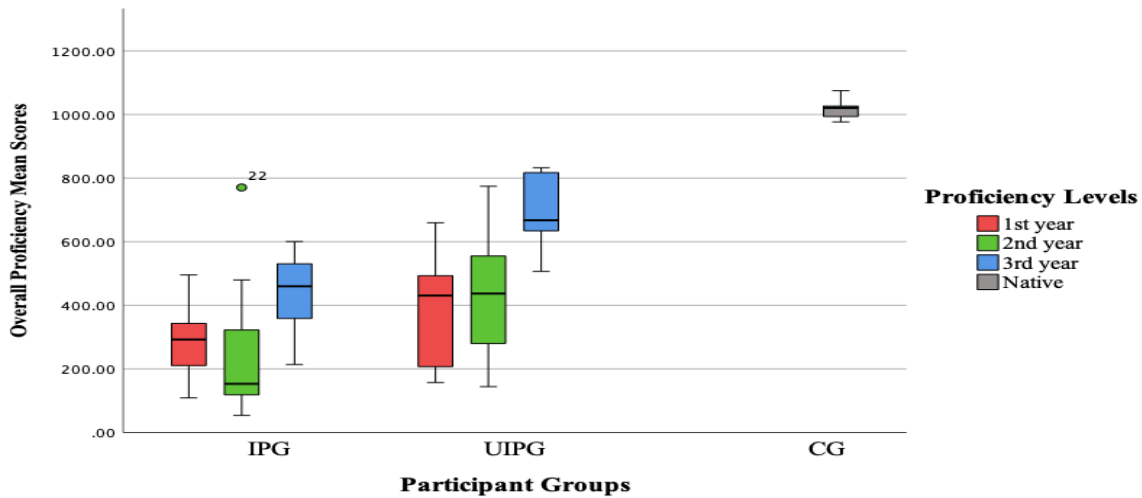
#### **4.5 Overall Proficiency**

To measure the overall proficiency of the participants, the percentage scores of the three proficiency constructs (i.e., fluency, accuracy, complexity) of each participant were added up. Table 4.28 summarizes the means of the IPG, UIPG, and CG on all the dependent variables examined, and it provides the means of the overall proficiency based on the added-up scores of the dependent variables (see also Figure 4.11). It is evident that the overall proficiency means of the UIPG were higher than those of the IPG across all proficiency levels. In addition, within each L2 group, the overall proficiency maintained a steady status from the first-year to second-year levels before it exhibited a considerable increase at the third-year level. Notwithstanding this observation, the overall proficiency of the third-year IPG participants was comparable to that of the first-year and second-year UIPG participants. Moreover, the control participants were irrefutably more proficient than the participants of both L2 groups across all proficiency levels.

**Table 4.28** Overall Proficiency Means of the IPG, UIPG, and CG

Variable	IPG			UIPG			CG
	1st year (n = 9)	2 year (n = 11)	3 year (n = 8)	1st year (n = 10)	2nd year (n = 10)	3rd year (n = 9)	Native (n = 9)
<b>Fluency</b>							
Past T Task	20.44	22.73	25.88	17.10	20.00	30.00	108.78
Present T Task	22.67	26.91	29.38	19.00	22.90	32.67	118.22
<b>Accuracy</b>							
Past T (3.s.m)	15.39	17.33	75.98	53.22	45.47	88.40	97.92
Past T (3.s.f)	9.72	13.38	50.75	46.40	50.96	70.19	98.68
Present T (3.s.m)	54.76	40.95	62.75	55.06	57.61	86.37	97.66
Present T (3.s.f)	56.72	43.06	57.60	47.36	57.50	84.19	100
Past Negation	5.56	15.45	14.79	44.09	31.00	48.82	100
Present Negation	58.28	37.07	70.83	65.68	67.14	96.30	100
<b>Complexity</b>							
Subjunctive Sub	21.83	19.15	35.90	27.45	41.18	77.27	100
Cause–Effect Sub	17.37	12.50	15.30	33.46	35.71	70.16	100
<b>Overall Proficiency</b>	282.73	248.53	439.14	408.82	429.47	684.36	1021.26

**Figure 4.11** Boxplot for the Overall Proficiency Means of the IPG, UIPG, and CG



A full factorial two-way ANOVA test was performed to examine the main effects of the two independent variables (the L2 participant group and L2 proficiency level), as well as their interaction effects, on the overall proficiency mean scores of the two L2 groups across the three

proficiency levels. The main effect of the L2 participant group factor was statistically significant ( $F(1,51) = 17.246, p < .001$ ) and yielded an effect size of .253, indicating that 25.3% of the variances in the L2 participants' overall proficiency mean scores were explained by the participant groups to which they belonged. The main effect of the proficiency level factor was likewise significant ( $F(2,51) = 10.349, p < .001$ ) and yielded an effect size of .289, indicating that 28.9% of the variances in the L2 participants' mean scores were explained by their proficiency levels. However, the interaction effect was not significant ( $F(2,51) = .573, p = .568$ , partial  $\eta^2 = .022$ ), indicating that the main effect of the L2 participant group factor did not depend on the L2 proficiency level factor or vice versa. Tukey's post hoc analyses were carried out to reveal the nature of the differences between the proficiency levels (Table 1k: Appendix K). The results showed that the first-year and second-year participants did not differ significantly from one another ( $p = .960$ ). However, the third-year participants differed significantly from both the first-year participants ( $p < .001$ ) and second-year participants ( $p < .001$ ). Based on the follow-up pairwise comparisons between the IPG and UIPG at each proficiency level (Table 2k: Appendix K), the L2 participant group factor was significant due to the significant outperformance of the UIPG over the IPG at the second-year level ( $p = .016$ ) and third-year level ( $p = .004$ ). In addition, based on the follow-up pairwise comparisons between all proficiency levels within each L2 group (Table 3k: Appendix K), the L2 proficiency level factor yielded a significant main effect because the third-year IPG participants were significantly more proficient than the second-year IPG participants only ( $p = .017$ ) and because the third-year UIPG participants were significantly more proficient than both the first-year UIPG participants ( $p < .001$ ) and second-year UIPG participants ( $p = .002$ ).

To examine whether the control participants differed from the L2 participants representing each proficiency level within each L2 group in their overall proficiency, six independent sample *t*-tests were performed. The results showed that the CG was significantly more proficient than both the IPG and UIPG across all proficiency levels (Table 4.29).

**Table 4.29** Result of the *t*-Tests on the Overall Proficiency Mean Scores of the CG and Each Proficiency Level Within Each L2 Group

Variable (I)	Variable (J)	<i>df</i>	<i>t</i>	Sig. <i>p</i>	95% Confidence Interval	
					Lower Bound	Upper Bound
CG	IPG					
Native ( <i>n</i> = 9)	1st year ( <i>n</i> = 9)*	9.308	-18.402	<.001	-828.85	-648.19
	2nd year ( <i>n</i> = 11)*	10.573	-11.677	<.001	-919.101	-626.36
	3rd year ( <i>n</i> = 8)*	7.857	-12.632	<.001	-688.73	-475.51
CG	UIPG					
Native ( <i>n</i> = 9)	1st year ( <i>n</i> = 10)*	9.687	-10.635	<.001	-741.32	-483.56
	2nd year ( <i>n</i> = 10)*	9.612	-9.718	<.001	-728.22	-455.35
	3rd year ( <i>n</i> = 9)*	9.213	-8.103	<.001	-430.62	-243.18

*Note.* \* = equality of variances was not assumed via Levene's *F* test (*p* < .05).

#### 4.6 Summary

This section provides an overall summary of the results revealed by the production data of the participants on the three proficiency constructs examined, namely fluency, accuracy, and complexity—which collectively constitute the participants' overall proficiency. The fluency of the two L2 groups (IPG and UIPG) was statistically comparable at each proficiency level (first year, second year, and third year) in both the past and present narrative tasks. Although the L2 participant group factor did not yield any significant main effects for one group over the other, the L2 proficiency level factor corresponded to significant main effects. This was mainly due to the outperformance of the third-year UIPG participants, who produced W/M at a significantly higher speed than both the first-year and second-year UIPG participants in both narrative tasks.

This finding can be argued to provide an advantage for the UIPG participants only since their fluency increased overtime (at the third-year level), contrary to their IPG counterparts, whose fluency did not exhibit any significant improvement as they advanced from the first-year to second-year to third-year levels. Nevertheless, the fluency of both L2 groups across all proficiency levels was significantly lower than that of the CG in both tense contexts.

With respect to the accuracy construct, the analysis of the past tense data in the 3.s.m and 3.s.f contexts yielded significant main effects for the L2 participant group factor, which was mainly due to the outperformance of the first-year and second-year UIPG participants over their IPG counterparts. The analysis likewise revealed significant main effects for the L2 proficiency level factor, which was due to the outperformance of the third-year participants over the first-year and second-year participants within both the IPG and UIPG in the 3.s.m context, and which was due to the outperformance of the third-year IPG participants over the first-year and second-year IPG participants in the 3.s.f context. The latter result may suggest that, in the 3.s.f context, only the accuracy of the IPG participants improved as their proficiency increased, unlike their UIPG counterparts, whose accuracy did not exhibit any appreciable improvement overtime. However, the medians of the interquartile accuracy range of the first-year and second-year UIPG participants were even higher than that of the third-year IPG participants. This observation indicates that the acquisition of the past tense with the 3.s.f form took place readily and rapidly at an earlier stage among the UIPG participants, compared to their IPG counterparts. Moreover, only the third-year UIPG participants were statistically on par with the CG in the 3.s.m context; however, both the IPG and UIPG across all proficiency levels significantly underperformed the CG in the 3.s.f context.

As for the accuracy of the participants on the present tense in the 3.s.m and 3.s.f contexts, the analysis revealed likewise an advantage for the UIPG over the IPG in the long run. That is, the accuracy mean scores of the third-year UIPG participants were significantly higher than that of their IPG counterparts in the 3.s.m and 3.s.f contexts. The advantage of the UIPG was further evidenced by the significant outperformance of the third-year UIPG participants over both the first-year and second-year UIPG participants in both agreement contexts, unlike the IPG, who did not exhibit any statistically significant improvement from the first-year to third-year levels in both agreement contexts. Moreover, only the UIPG participants were statistically on par with the CG in their accuracy on the present tense in the 3.s.m context; however, the CG significantly outperformed both L2 groups across all proficiency levels in the 3.s.f context.

Regarding accuracy on negation, the analysis of the past negation data revealed a notable advantage for the UIPG over the IPG, particularly at the first-year and third-year levels. Remarkably, the first-year UIPG participants (beginner) were significantly more accurate than the third-year IPG participants (advanced)—debatably suggesting that fossilization might have taken place at an early stage among the IPG participants. However, the accuracy of both L2 groups across all proficiency levels was significantly lower than that of the CG.

As for the analysis of the verbal present negation data, it yielded likewise a clear advantage for the UIPG over the IPG because of the significant and near-significant outperformance of the second-year and third-year UIPG participants over their IPG counterparts, respectively. The advantage of the UIPG was further supported by the fact that the third-year UIPG participants were significantly more accurate than both the first-year and second-year UIPG participants, whereas the third-year IPG participants did not differ significantly from the first-year IPG participants. Moreover, the accuracy of the CG on the verbal present negation was

significantly higher than that of both L2 groups across all proficiency levels, except for the third-year UIPG participants. This, in turn, implies that only the UIPG participants were gradually able to achieve native proficiency in the their acquisition of this structure.

Regarding complexity, the analysis of the error-free production data of the subjunctive subordination and cause–effect subordinations revealed again a significant advantage for the UIPG over the IPG at the second-year and third-year levels. The appreciable advantage of the former group was further evidenced by the significant outperformance of the third-year UIPG participants over both the first-year and second-year UIPG participants, unlike the IPG, whose error-free production of these two complex structures did not exhibit any significant improvement as proficiency increased. However, the mean scores of both L2 groups across all proficiency levels were significantly lower than those of the CG on both structures.

The statistical analysis of the added mean scores of the participants on every proficiency construct revealed that, although the IPG and UIPG were comparable in their overall proficiency at the first-year level, the former was significantly less proficient than the latter at the second-year and third-year levels. Additionally, the overall proficiency of the UIPG exhibited a steady improvement as proficiency increased (precisely at the third-year level), whereas the overall proficiency of the IPG was statistically flat across all proficiency levels. However, both L2 groups across all proficiency levels were significantly less proficient than the CG.

The above findings, which unequivocally revealed a conclusive advantage for the UIPG over the IPG, have theoretical implications for three factors that play a paramount role in SLA: input, L1 transfer, and implicit versus explicit learning. The findings also bear practical and applied implications for Arabic L2 teachers, textbook writers, and researchers. Detailed discussions of both types of implications are provided below in Chapter 5.

## **Chapter 5 Discussion and Conclusion**

### **5.1 Overview**

The current chapter aims to discuss the results reported in Chapter 4. The discussion starts by addressing the first research (RQ) pertinent to the nature of the language output of the integrated program group (IPG) and that of the unintegrated program group (UIPG). Subsequently, the discussion focuses on interpreting the similar and different performance of both L2 groups in light of three key factors that contribute the most to the development of SLA: input, L1 transfer, and implicit versus explicit learning. Lastly, the chapter concludes by highlighting the pedagogical implications of the study and by acknowledging its limitations.

### **5.2 Discussion of the Research Questions**

This section is grouped into two subsections. The first offers a concise discussion of the first RQ related to the fluency, accuracy, complexity, and overall proficiency of the IPG participants and their UIPG counterparts. The main aim of the first subsection is to provide preliminary explanations for the similarities and differences in the performance of all participants and to lay the basis for the second subsection, which aims at addressing the second RQ concerned with the role of input, L1 transfer, and implicit versus explicit learning in the performance of the L2 participants on the target structures.



## **5.2.1 RQ1: Nature of the language output of the IPG and UIPG**

### ***5.2.1.1 Fluency***

This section discusses the findings pertaining to the first subquestion of RQ1, namely: are there any similarities and differences between the IPG and UIPG in terms of their fluency when narrating past and present events? Recall, the fluency of the participants was measured by counting words produced per minute (W/M). The results revealed that both L2 groups narrated past and present events at a statistically comparable speed at each proficiency level.

However, when examining the improvement of fluency from the first-year to second-year to third-year levels within each L2 group, only the UIPG participants improved significantly over time. That is, the third-year UIPG participants produced W/M at a significantly higher speed than the first-year and second-year UIPG participants in both the past and present tense narrative tasks, whereas the IPG participants produced W/M at a statistically comparable speed across all levels in both tasks. Only the data of the UIPG are congruent with the ACTFL guidelines (2012), which postulate that speakers exhibit a tendency to narrate and describe events in all time frames as their proficiency increases. This, in turn, can be argued to provide an advantage for the UIPG over the IPG because the fluency of the former improved significantly over time, whereas the fluency of the latter was statistically flat.

When comparing the W/M produced by the IPG and UIPG to those produced by the control group (CG), the CG was significantly more fluent than both L2 groups across all proficiency levels when narrating past events and present events alike. In fact, the W/M produced by the control participants were approximately three times higher than those produced by the L2 participants (see Table 4.1, Figures 4.1 and 4.2: Chapter 4). This finding aligns well with those reported in previous studies that found the following: (a) W/M produced by Arabic

native speakers were significantly higher than those produced by Arabic heritage speakers (Albirini, 2018); (b) oral and written English W/M produced by English L1 speakers were significantly higher than those produced by Arabic L1 learners of English as an L2 (Albirini, 2019); (c) written Arabic W/M produced by native speakers of Arabic were significantly higher than those produced by English L1 advanced L2ers of Arabic (Raish, 2018).

However, a question remains unanswered, namely: why did not the fluency of the IPG improve significantly as proficiency increased? This might be explained in view of the assumption that L2ers become more fluent when focus on function is prioritized and focus on form is neglected (R. Ellis & Barkhuizen, 2005). Given that the simultaneous exposure to two Arabic varieties from the outset of the learning process violates the clarity of input principle (Alhawary, 2013, 2021) and is considered a main source of confusion for L2ers (as reported in Al-Batal & Glakes, 2018; Huntley, 2018), receiving input in MSA and EA perhaps led the IPG participants to cognitively concentrate on producing accurate and authentic rather than fluent language output. Possibly, this made them hesitant and undecided about, for example, which variety they would use in their description of the two narrative tasks.<sup>9</sup> As a consequence, they probably attempted to keep their narration to the minimal level required, which in turn prevented them from implementing communicative strategies (e.g., circumlocution, paraphrasing, and illustration) that speakers usually implement to extend the length of their speech (ACTFL, 2012; Fakhri, 1984). If this assumption is accepted, it can be further argued that receiving input in one variety only (à la the clarity of input principle) assisted the UIPG participants to retrieve the Arabic vocabulary repertoire and access the morphosyntactic target structures more easily and

---

<sup>9</sup> This argument is supported by the fact that, during the elicitation sessions, some of the IPG participants were reluctant about which variety they should use, evident in their question to the researcher whether they should use MSA or EA (see Section 3.5: Chapter 3).

efficiently over time (at the third-year level), compared to their IPG counterparts, whose exposure to two varieties had a negative impact on their fluency.

### ***5.2.1.2 Accuracy***

This section discusses the findings pertaining to the second subquestion of RQ1, namely: are there any similarities and differences between the IPG and UIPG in terms of their accuracy on the target structures of past tense and present tense, as well as past negation and verbal present negation? The overall findings revealed an advantage for the UIPG participants over their IPG counterparts in their accuracy on these structures, as will be discussed further below.

#### *5.2.1.2.1 Past tense*

Given that MSA and EA converge in the past tense structure (see Section 3.3.1.1: Chapter 3) and that both the IPG and UIPG participants were comparable in terms of their L1, prior exposure to Arabic, and proficiency-placement levels (see Section 3.2.2.2: Chapter 3), it would be expected that both L2 groups would perform comparably on this structure across all proficiency levels. However, this was not the case because the UIPG participants were significantly more accurate than their IPG participants at the first-year and second-year levels in their accuracy on the past tense with both the 3.s.m and 3.s.f agreement forms. In fact, the accuracy mean scores of the first-year and second-year IPG participants were below the 20% accuracy level in both agreement contexts, compared to their UIPG counterparts, whose accuracy mean scores were almost at the 50% accuracy level in both agreement contexts (see Table 4.4: Chapter 4). It was also visually evident that the interquartile accuracy ranges of the first-year and second-year IPG participants on both agreement forms were clustered at the bottom of the two boxplots, compared to their UIPG counterparts, whose interquartile accuracy ranges exhibited greater variability (see Figures 4.3 and 4.4: Chapter 4). The bottom clustered accuracies of the first-year and second-

year IPG participants may indicate that acquisition of the past tense inflected for the 3.s.m and 3.s.f forms did not take place among them, or (at best) it was barely merging in their interlanguage systems. In contrast, the widespread accuracies of their UIPG counterparts signify that their acquisition of this structure was in progress.

However, the IPG improved significantly in their accuracy on the past tense over time—at the third-year level. Nevertheless, their interquartile accuracy ranges on both agreement forms overlapped with those of the first-year and second-year UIPG participants (see Figures 4.3 and 4.4: Chapter 4). Thus, it is plausible to endorse the argument that the development of the past tense with both agreement forms took place rapidly at an earlier stage among the UIPG participants, compared to their IPG counterparts, whose production of this structure became more accurate after almost 3 years of learning. A crucial question that has not been addressed yet is why the IPG participants experienced a delay in their acquisition of this structure. One apparent cause can be attributed to the input frequency and presentation timing of the past tense in their instructional textbook, as will be discussed in Section 5.2.2.1 below.

When comparing the accuracy of the IPG and UIPG to that of the CG on the past tense, it was found that the CG significantly outperformed both the IPG and UIPG across all levels, except for the third-year UIPG participants in the 3.s.m context only. This finding indicates that, when it comes to the acquisition of this target structure, only the UIPG participants were gradually able to approximate further toward native proficiency over time. However, this is not to say the third-year UIPG participants achieved complete native competence in this structure because they differed significantly from the control participants in the 3.s.f context.

One important finding revealed by the fine-grained qualitative analysis of the incorrect rule application of the past tense structure is that the error that was due to producing present

tense forms instead of their past counterparts was prevalent among the first-year and second-year participants of both the IPG and UIPG, but it was rarely produced by the third-year participants of both L2 groups (see Tables 4.5 and 4.6: Chapter 4). The considerable decline of this error at the third-year level of both L2 groups highlights the superior advantage of input frequency over L1 transfer, as will be discussed in Section 5.2.2.1 below.

Another finding is that both L2 groups encountered some difficulties in processing and producing the past tense with the 3.s.f agreement form. This is strongly supported by the fact that, while the agreement error was the least frequent error in the 3.s.m context, it was the predominantly produced error in the 3.s.f context and was prevalently persistent across all proficiency levels of both L2 groups (see Tables 4.5 and 4.6: Chapter 4). The implication of this finding is that the L1 transfer effect (from English) was afoot and was not minimized by the frequency effect of the past tense inflected for 3.s.f in the input, as will be discussed in Section 5.2.2.2. Moreover, the lack of diacritics in the input, as a visual (explicit) enhancer, was perhaps another reason for the persistence of this error, as will be discussed in Section 5.2.2.3.

#### *5.2.1.2.2 Present tense*

The analysis of the L2 present tense data revealed that the participants of both the IPG and UIPG exhibited a higher accuracy rate in their use of the present tense than the past tense in both agreement contexts. A reason for this asymmetrical accuracy performance on these tense structures can be attributed to the frequency and presentation timing of these two structures in the instructional textbooks used by the participants, as will be discussed in Section 5.2.2.1.

With regard to the performance of the IPG and UIPG on the present tense, the quantitative analysis revealed that the UIPG significantly outperformed the IPG in the long run, particularly at the third-year level, in both the 3.s.m and 3.s.f agreement contexts. The advantage

of the UIPG is further supported by the within-group analysis. Although the third-year UIPG participants significantly outperformed both the first-year and second-year UIPG participants in both contexts, the third-year IPG participants did not significantly differ from the first-year IPG participants in the 3.s.m context and did not significantly differ from both the first-year and second-year IPG participants in the 3.s.f context. The apparent reason for this significantly asymmetric performance between the two L2 groups is likely attributed to the simultaneous exposure to MSA and EA by the IPG participants, which perhaps obstructed their comprehension of the input, as will be discussed in Section 5.2.2.1 below.

When comparing the accuracy of the IPG and UIPG to that of the CG on the present tense, the CG significantly outperformed both L2 groups across all levels, except for the third-year UIPG participants in the 3.s.m context only. This finding indicates that, when it comes to the acquisition of this target structure, only the UIPG participants were eventually able to approximate further to native-like proficiency. However, this is again not to say the third-year UIPG participants achieved ultimate native competence in this structure because they differed significantly from the CG in the 3.s.f context, which highlights the critical effect of L1 transfer in the acquisition of this structure, as will be discussed further in Section 5.2.2.2 below.

An interesting finding revealed by the qualitative analysis of the present tense data is that both L2 groups exhibited a tendency to converge in their use of the default form (3.s.m) in their production of the present tense in the 3.s.f context. This findings supports previous research that found a similar pattern among Arabic L2ers from various L1 backgrounds, including English and French (Alhawary, 2005, 2009b); English, French, Japanese, and Spanish (Alhawary, 2009a); and Chinese and Russian (Alhawary, 2019). However, it is important to note that, while this agreement error decreased as proficiency increased within the UIPG (from the first-year to

second-year to third-year levels), it was prevalent in the data of the IPG and did not show a decreasing trend over time (see Tables 4.10 and 4.11: Chapter 4). The role of L1 transfer can account for the asymmetrical behavior of this error in the data of the IPG and UIPG, as will be discussed in Section 5.2.2.2 below.

#### *5.2.1.2.3 Past negation*

Although the IPG and UIPG received a similar amount of Arabic input and came from the same L1 background, the latter significantly outperformed the former at the first-year and third-year levels in their accuracy on the past negation. However, the pairwise comparison between the three proficiency levels within each L2 group did not yield any significant differences between all proficiency levels within either L2 group. The significant advantage of the UIPG over the IPG, as well as the nonsignificant improvement from the first-year to second-year to third-year levels within both L2 groups, is argued to be the ultimate outcome of the linguistic dissonance, resulting from the negation optionality in the input received by the participants of both L2 groups, as will be discussed in detail in Section 5.2.2.1 below.

The finding concerned with the nonsignificant improvement from the first-year to second-year to third-year levels within either one of the L2 groups is inconsistent with that reported in Albirini et al.'s (2019) study, which found the accuracy of English L1 advanced L2ers of Arabic (50.76%) to be significantly higher than that of their beginner counterparts (11.33%) on the past negation. The inconsistency here might be attributed to the nature of tasks used in the current study and that used in Albirini et al.'s (2019) study. While the current study elicited this target structure spontaneously from the participants without drawing their attention to it, the task implemented in Albirini et al.'s (2019) study entailed transforming affirmative past sentences into their negative counterparts. It is possible that such a task raised the participants'

awareness, promoting the advanced participants to rely on their metalinguistic knowledge and produce this structure with greater caution and accuracy.

When comparing the accuracy of the IPG and UIPG to that of the CG on the past negation, the control participants outperformed the L2 participants of both L2 groups across all proficiency levels. This result conforms to that reported in Albirini et al.'s (2019) study, which found the accuracy of native Arabic speakers on the past negation to be significantly better (98.04%) than English L1 learners of Arabic as an L2 at the beginner and advanced levels (11.33% and 50.76%, respectively). This strongly highlights that the Arabic past negation system poses a difficulty for Arabic L2ers, irrespective of their proficiency levels. The only obvious reason for this difficulty is again attributed to the simultaneous learning of the two MSA negation constructions (*mā + perfect verb* and *lam + imperfect verb*) by the UIPG, in addition to the EA negation system (*mā + perfect verb + š*) by the IPG. Thus, further pedagogical intervention is much needed to facilitate the acquisition of the Arabic past tense negation structures by English L1 speakers (see Section 5.3 below).

An important finding revealed by the qualitative analysis is that both L2 groups exhibited a tendency to use the *mā + perfect verb* construction more often than the *lam + imperfect verb* construction. Precisely, of the 153 correct MSA tokens produced by both L2 groups, only 16 tokens exhibited the latter construction (10.5%), whereas the remaining 137 tokens exhibited the former construction (89.5%). This finding is consistent with those reported in Albirini et al. (2019) and Alhawary's (2009a) studies, which found a similar tendency among English L1 learners of Arabic as an L2. Such a finding can be explained considering the presentation timing of these constructions in the input, as will be discussed in Section 5.2.2.1.



#### *5.2.1.2.4 Verbal present negation*

Contrary to what was revealed by the past negation data, the verbal present negation data showed that the IPG and UIPG exhibited a tendency to produce this structure (in MSA) at a high accuracy rate almost across all proficiency levels. This finding suggests that the verbal present negation emerged in the L2 participants' interlanguage systems more rapidly and readily than its past counterpart. This could be attributed to the frequency and presentation timing of both negation structures in the participants' textbooks, as will be discussed in Section 5.2.2.1. This finding also adds support to previous research that found a similar pattern of accuracy on both negation structures among English L1 beginner L2ers of Arabic (Albirini et al., 2019; Alhawary 2009a) and English L1 advanced L2ers of Arabic (Albirini et al., 2019).

Nevertheless, the UIPG exhibited again a higher accuracy rate than that of the IPG on the verbal present negation due to the significant and near-significant outperformance of the first-year and third-year UIPG participants over their IPG counterparts. The advantage of the UIPG is further supported by two pieces of evidence: First, the third-year UIPG participants differed significantly from the first-year and second-year UIPG participants, whereas the accuracy of the third-year IPG participants was statistically on par with that of the first-year IPG participants. Second, although the CG significantly outperformed the IPG group across all proficiency levels, the accuracy mean score of the CG was statistically comparable to that of the third-year UIPG participants. Only the finding revealed by the UIPG data is consistent with that reported in Albirini et al.'s (2019) study, which found the performance of English L1 advanced L2ers of Arabic to be significantly better (84.85%) than that of their beginner counterparts (57.33%) but statistically comparable with that of native speakers of Arabic (100%).

The asymmetrical performance between the IPG and UIPG on the present negation can possibly be attributed to the confusion experienced by the IPG participants, who were exposed simultaneously to the MSA system (*lā + imperfect verb*) and its EA counterpart (*ma + bi + imperfect verb + š*). Likely, this optionality in the input caused confusion to the IPG and resulted in problematic production, as will be discussed further in Section 5.2.2.1 below. In addition, lack of explicit instruction of the EA negation system, which involves a complex rule, was possibly another cause for such a finding, as will be discussed in Section 5.2.2.3.

### **5.2.1.3 Complexity**

This section discusses the findings pertaining to the third subquestion of RQ1, namely: are there any similarities and differences between the IPG and UIPG in terms of their error-free production of the target complex structures of subjunctive subordination and cause–effect subordination? The quantitative analysis of the participants’ error-free production of the subjunctive and cause–effect subordinations revealed that the UIPG significantly outperformed the IPG at the second-year and third-year levels on both structures. The advantage of the former was further evidenced by the fact the third-year UIPG participants significantly outperformed the first-year and second-year UIPG participants, whereas the performance of the IPG was statistically flat across all proficiency levels. This can explain why there was an interaction between the two independent variables of the L2 participant group and L2 proficiency level. The significant improvement in the error-free production of these two complex structures over time was dependent on the programs to which the L2 participants belonged. In other words, only those of the UIPG improved significantly in their error-free production of these complex structures as their proficiency increased (see Figures 4.9 and 4.10: Chapter 4).

When comparing the performance of the IPG and UIPG to that of the CG on the two subordination structures, the CG significantly outperformed the IPG and UIPG across all proficiency levels. This finding is congruent with previous research that found the following: (a) Arabic native speakers produced significantly more complex language than heritage speakers of Arabic (Albirini, 2018); (b) English native speakers produced oral and written complex language at a significantly higher rate than Arabic L1 learners of English as an L2 (Albirini, 2019). However, the findings of the current study do not align with those reported in Raish's (2018) study, which measured the complexity construct in writing by calculating words produced per T-unit (W/T), which were statistically comparable between native Arabic speakers and English L1 advanced L2ers of Arabic (in the persuasive task only). The inconsistencies of these findings is justified due to the nonexistence of a uniform definition and measure of complexity, which resulted in asymmetrical complexity findings across SLA studies (e.g., Housen et al., 2012). To elaborate, Albirini (2018, 2019) measured complexity quantitatively only by calculating the ratio of complex T-units (e.g., subordination) to the total number of T-units produced (i.e., all sentences), regardless of the correct and incorrect production of those T-units. Raish (2018) likewise measured complexity quantitatively but by counting words produced per T-unit (W/T), regardless of the correct and incorrect production. However, the current study avoided this limitation by eliciting the complex T-units (i.e., subjunctive and cause-effect subordinations) in obligatory contexts only and by implementing qualitative and quantitative analyses, entailing ratio calculation of the error-free production only of these two complex T-units to the total number of these complex T-units. It was assumed that adopting this micro-based measure would provide more insight into the development of these complex structures over time in the L2ers' interlanguage systems and would shed light on some areas that might pose difficulties to them.

Given the great homogeneity between the IPG and UIPG participants in terms of their L1, prior exposure to Arabic, and proficiency-placement levels, they would be expected to perform comparably on the subjunctive and cause–effect subordination structures. However, this was not the outcome, and this in turn leaves room for accounting for this asymmetrical performance between the two L2 groups in light of the input factor. That is, the obscurity in the input the IPG participants received (i.e., exposure to MSA and EA subordination structures) is most likely the prime cause for their flat performance across all proficiency levels, as will be discussed in Section 5.2.2.1 below.

Other important findings revealed by the qualitative analyses are pertinent to the use of the simple verbal noun structure instead of complex subjunctive subordination structure and the use of the simple *ʔiḍāfah* structure instead of the complex cause–effect subordination structure. Recall, the subjunctive subordination can be optionally substituted by a verbal noun derived from the subjunctive subordinate verb. Likewise, the complex cause–effect subordination headed by the conjunctions *liʔann* or *ʕašān/ʕalašān* “because” can be optionally substituted by an *ʔiḍāfah* phrase, headed by the MSA phrase *bisabab* or the EA conjunction *ʕašān* “because of” and followed by a genitive noun phrase. These two optional structures are considered simple rather than complex, so they were not counted toward the total production of the two complex subordination structures. However, they were analyzed separately to examine the extent to which the L2 participants of both L2 groups across all proficiency levels opted for the simple structure in comparison to its complex counterpart. The analysis revealed that the first-year participants of both L2 groups opted for the simple structures the most, then the second-year participants, and then the third-year participants (see Tables 4.21 and 4.25: Chapter 4). The indication of this finding is that the beginner and intermediate participants exhibited a tendency to avoid the

complex subjunctive and cause–effect subordination structures, which possibly added extra-cognitive loads on them. However, this was not an issue for the advanced L2ers, who produced the complex structures more often than those at the beginner and intermediate proficiency levels. This result is consistent with the ACTFL proficiency guidelines (2012), which suggest that advanced speakers exhibit abundant language that enables them to generate lengthy and complex output, unlike the intermediate and beginner L2ers. In addition, this finding indicates that the L2 participants were willing to take risks by using a wide range of different complex structures (R. Ellis & Barkhuizen, 2005; Norris & Ortega, 2009) as their proficiency increased. However, it is worthy to note that the error that was due to lack of suppliance of the subjunctive subordination (see Table 4.22: Chapter 4) and cause–effect subordination (see Table 4.26: Chapter 4) was persistent across all proficiency levels in the IPG, whereas it considerably decreased at the third-year level in the UIPG. Thus, this finding might suggest that only the UIPG participants were able to produce more complex language as their proficiency increased, particularly at the third-year level, and this finding adds support to the fluency results discussed earlier.

#### ***5.2.1.4 Overall proficiency***

This section discusses the findings pertaining to the fourth subquestion of RQ1, namely: are there any similarities and differences between the IPG and UIPG in terms of their overall proficiency?

Recall, the overall proficiency was measured by the aggregate calculation of the scores attained in the three proficiency constructs: fluency, accuracy, and complexity (see Table 4.28: Chapter 4). The statistical analysis revealed a significant advantage of the UIPG over the IPG at the second-year and third-year levels. Additionally, the third-year UIPG participants were significantly more proficient than the first-year and second-year UIPG participants. In contrast,

the overall proficiency of the third-year IPG participants was significantly more than that of the second-year IPG participants but statistically comparable with that of the first-year IPG participants. Based on the assumption that “without understanding the language, no learning can occur” (Gass & Selinker, 2008, p. 305; see also Alhawary, 2013, 2021; Corder, 1967; Gass & Mackey, 2002; Krashen, 1985; Long, 1996; Loschky, 1994; Nation & Macalister, 2010; Sachs et al., 1981; Snow et al., 1976) and that optionality in the input might hinder comprehension (Alhawary, 2013, 2021; Eisele, 2018; Nation & Macalister, 2010; Schwartz & Causarano, 2007), the strong indication of the current findings is that exposure to one Arabic variety (i.e., input transparency) facilitated comprehension among the UIPG participants, which in turn led them to make significant gains in their learning process over time. In contrast, exposure to two Arabic varieties (i.e., input obscurity) prevented the IPG participants from comprehending the input, which consequently hindered them from making significant gains—despite the prolonged years of learning Arabic as an L2. Interestingly, the overall proficiency mean score of the third-year IPG participants was quite similar to those of the first-year and second year UIPG participants. This notable similarity may debatably suggest that the Arabic learning gains achieved by the IPG over a span of 3 years appear to be comparable with those achieved by the UIPG in 1–2 years.

When comparing the overall proficiency of the IPG and UIPG to that of the CG, it was revealed that the control participants were significantly more proficient than the L2 participants of the IPG and UIPG across all proficiency levels. This finding was expected, given that Arabic L2ers can barely develop a “basic competence” (Abboud, 1968) and “working knowledge” (Alhawary, 2013, 2021) during the first 2–3 years (i.e., the normal length of learning Arabic as an L2 in U.S. institutions). This finding also contradicts those reported in Leddy-Cecere (2018)

and Nassif's (2018) studies, which found the simultaneous learning of MSA and CA to be beneficial to Arabic L2ers in terms of developing native-like linguistic competence.

Moreover, an important issue strongly relevant to the current discussion pertains to the question of whether the simultaneous exposure to MSA and EA equipped the IPG participants with sociolinguistic (diglossic) competence similar to that of Arabic native speakers, as claimed by advocates of the IA (Al-Batal, 1992, 1995, 2018; Leddy-Cecere, 2018; Nassif, 2018; Shiri & Joukhadar, 2018; Younes, 1990, 2015) and as demanded by many survey-based studies that found a positive attitude held by Arabic L2ers toward such an instructional practice (Al-Batal & Glakes, 2018; Belanp, 2006; Hashem-Aramouni, 2011; Huntley, 2018; Husseinali, 2006; Isleem, 2018; Kuntz & Benlap, 2001; Palmer, 2007, 2008; Shiri, 2013). This claim was evaluated based on comparing the data of the IPG (i.e., Arabic L2ers of MSA and EA) to that of the CG (i.e., native speakers of EA). It was found that the control participants opted for the EA variety all the time in their production. Such a finding would be anticipated due to the simple nature of the implemented picture-based tasks, which were about simple daily-life topics that entail utilizing the L variety (CA) rather than the H variety (MSA; Ferguson, 1959a). In contrast, the L2 IPG participants opted predominantly for MSA rather than EA in their production. This is so despite the fact their responses on the question related to the frequency of code-switching behavior in the LHQ distributed to all L2 participants revealed that they would code-switch between MSA and EA at a significantly higher rate of frequency than their UIPG counterparts (see Section 3.2.2.2.5: Chapter 3). The following provides a snapshot of all MSA and EA tokens produced by the IPG participants in the target structures in which the two varieties diverge:

- Of the 457 present tense tokens produced, only 12 tokens were inflected correctly for the agreement and tense features in EA (0.03%). Only one of these 12 tokens was completely

correct, where the verb was produced with the habitual prefix *bi-* in the 3.s.m context. However, none were produced by the third-year participants.

- Of the 283 past negation tokens produced, only 58 tokens were in EA (20.5%): 11 by first-year participants, 37 by second-year participants, and 10 by third-year participants. Remarkably, only two of these EA tokens were completely correct and produced by one third-year participant in the 3.s.m context (0.01% of all past negation tokens).
- Of the 262 verbal present negation tokens produced, only 46 tokens were in the EA variety (17.6%): 10 by first-year participants, 32 by second-year participants, and 4 by one third-year participant. Remarkably, only two tokens were completely correct and were produced by one first-year participant with the verb *yihibb* “to like” in the 3.s.m context (0.01% of all verbal present negation tokens).
- Of the 313 subjunctive subordination tokens produced, only 6 tokens were considered completely correct in the EA variety (0.02%), and they were produced by two first-year participants and one second-year participant. Interestingly, the finite subordinate clauses in these tokens were not subcategorized by imperfective verbs in the matrix clause but rather by the active participles *ṣāyiz/a* “wanting,” *lāzim* “obligating,” and *mumkin* “could be” (e.g., *hiya ṣāyiz-a ti-sāfir* “she wants [to] travel”). This indicates that these participants had difficulty in producing the habitual and progressive marker *bi-* with imperfect verbs (i.e., the two contexts in which the questions of the subjunctive subordination task were asked).
- Of the 214 cause–effect subordination tokens produced, only 11 tokens were completely correct when using the EA conjunction *ṣašān/ṣalašān* “because” (0.05%): 5 by first-year participants, 3 by second-year participants, and 3 by third-year participants. Remarkably,



the subordinate clauses in these correct tokens were subcategorized by adjectives rather than verbs in the matrix clauses (e.g., *hiya bārid-a giddan ʕašān ʔat-ʔaqs bārid giddan* “she is very cold because the weather is very cold”).

Thus, the findings mentioned above clearly indicate that the participants learning MSA and EA (i.e., IPG) developed neither linguistic nor sociolinguistic competencies similar to those of the CG (i.e., native speakers of EA). The findings also show that the third-year participants adhered to the MSA variety more strictly than the first-year and second-year participants (cf. Leddy-Cecere, 2018; Nassif, 2018). In fact, these findings confirm the concerns raised by Alhawary (2021) regarding the data collection methods employed in Nassif’s (2018) study, as well as in Shiri and Joukhadar’s (2018) study, which mainly relied on presentations and dialogue skits. The increase in the code-switching behavior in the output of the participants who took part in these two studies is attributed to the possibility that they practiced and perhaps memorized by rote the content of their presentations and skits.

Moreover, the predominant use of MSA by the IPG participants can be well justified in light of the complex dynamic systems theory (CDST), which predicts that development of the L2 (sub)systems is restricted by the internal and external resources available to L2ers (de Bot et al., 2007; de Pot & Larsen-Freeman, 2011; Larsen-Freeman, 1997, 2017). In other words, these participants opted for the MSA variety because it is recycled more frequently in their instructional textbook, more similar to their L1, and explicitly and visually presented in the input, as will be discussed in the immediate section below.

### **5.2.2 RQ2: Factors affecting the performance of the IPG and UIPG**

This section is an attempt to provide detailed interpretations of the current findings by answering RQ2 regarding the factors affecting the performance of the IPG versus that of the UIPG.

Discussion of such factors is focused here on the role of input, L1 transfer, and implicit versus explicit learning in the participants' performance on the target structures. These factors play interconnected roles in the development of the target morphosyntactic structures examined in the current study: past tense, present tense, past negation, verbal present negation, subjunctive subordination, and cause–effect subordination.

However, before proceeding, it is crucial to illustrate the similarities and differences among English (i.e., the L2 participants' L1), MSA, and EA in the target structures by providing six typological pairings. Such pairings are important for examining the role of input and implicit versus explicit learning in facilitating and minimizing the L1 transfer effect in the acquisition of the target structures by the IPG and UIPG participants.

- **Past Tense:** English, MSA, and EA converge in this structure since all of them exhibit a past tense feature. Thus, the L2 participants would not be expected to encounter any difficulties in acquiring this tense feature from early on due to its presence in their L1. However, English structurally diverges from Arabic in the sense that the verbal agreement paradigm of the former does not exhibit any person, number, and gender agreement features in the past tense (e.g., Alhawary, 2005, 2009a, 2009b, 2019). Therefore, the L2 participants would likely go through some difficulties in acquiring the Arabic past gender agreement feature due to its absence in their L1.
- **Present Tense:** English, MSA, and EA converge in the present tense since all of them exhibit a present tense feature. However, English and MSA diverge from EA in that the habitual prefix *bi-* must be produced with the imperfect verb in EA. Thus, the L2 participants would not be expected to encounter any difficulties in acquiring the MSA present tense feature from early on due to its presence in English. However, the EA

structure would be expected to pose some difficulties for the IPG participants due to the absence of the EA habitual marker in their L1. Moreover, English is structurally different from Arabic in the sense that the verbal agreement paradigm of the former is impoverished. Although English exhibits partial verbal present agreement in person through the third-person singular suffix *-s* and zero suffix elsewhere, it exhibits neither number agreement nor gender agreement (e.g., Alhawary, 2005, 2009a, 2009b, 2019). Therefore, the L2 participants would be expected to encounter some difficulties in acquiring the verbal present gender agreement in Arabic due to its absence in their L1.

- **Past Negation:** English structurally diverges from MSA in the *mā + perfect verb* construction and from EA in the *ma + perfect verb + š* construction; however, English and MSA converge in the *lam + imperfect verb* construction. Thus, on the one hand, the L2 participants would be expected to encounter some difficulties in acquiring the former MSA construction and EA construction due to their absence in English. On the other hand, the latter MSA construction would be expected to be relatively easy to acquire due to its presence in English.
- **Verbal Present Negation:** English and MSA converge in the *lā + imperfect verb* construction, but English and MSA are structurally different from EA in the *ma + bi + perfect verb + š* construction. Thus, while it would be expected for the MSA construction to be easy to acquire from early on due to its presence in English, the EA construction would be expected to pose some difficulties for the IPG participants because of its complex nature and due to its absence in English.
- **Subjunctive Subordination:** English and MSA converge in the *verb + ʔan “to” + imperfect verb* construction, but English and MSA are structurally different from EA in

the *active participle/bi + verb + imperfect verb* construction. Thus, the MSA structure would be expected to be easy to acquire from early on due to its structural proximity in English and MSA (+ complementizer). However, the EA construction would be expected to pose some difficulties for the IPG participants due to the structural differences between English (+ complementizer) and EA (- complementizer).

- **Cause–Effect Subordination:** English, MSA, and EA converge in the construction of matrix clause + *liʔann* “because” + subordinate clause containing a nominal-verbal nominal-verbless sentence, but English and MSA are structurally different from EA in the construction of matrix clause + *ʕašān/ʕalašān* “because” + subordinate clause containing a nominal-verbal, nominal-verbless, or verbal sentence. Thus, the former construction would be expected to be relatively easy to acquire from early on due to its structural proximity in English, MSA, and EA. However, the latter construction would be expected to pose some difficulties for the IPG participants because the conjunction *ʕašān/ʕalašān* “because” can also be followed by a verbal sentence (i.e., VSO).

### 5.2.2.1 Input

The forgoing discussion in the current and preceding chapters showed that the UIPG significantly outperformed the IPG in the acquisition of the target morphosyntactic structures in which MSA and EA diverge: present tense, past negation, verbal present negation, subjunctive subordination, and cause–effect subordination. In addition, the UIPG participants improved significantly over time, particularly at the third-year level, in their acquisition of these target structures, except for the past negation. In contrast, the IPG participants did not exhibit any significant improvement, particularly from the first-year to third-year levels, in their acquisition of all these structures. Two critical questions need to be addressed here, namely: (a) why did not

the IPG participants make significant learning gains identical to those of their UIPG counterparts in their acquisition of the target structures in which MSA and EA diverge? and (b) why did not the UIPG participants exhibit significant improvement over time in their acquisition of the MSA past negation system, as they did in other target structures?

A rational and plausible answer to the two questions mentioned above revolves around the role of input comprehension, which is considered an essential factor that “does set the scene for learning to take place” (Gass & Selinker, 2008, p. 305). It is a given that if the *input* was not comprehended, this means it did not turn into *intake* (Alhawary, 2013, 2021; Corder, 1967; Gass & Selinker, 2008; Krashen, 1985; Long, 1996; Loschky, 1994; Nation & Macalister, 2010; Sachs et al., 1981; Snow et al., 1976). Based on what was reported in previous studies, which found simultaneous exposure to two Arabic varieties (i.e., optionality in the input) to be confusing to Arabic L2ers (e.g., Al-Batal & Glakes, 2018; Huntley, 2018), it is plausible to argue that the simultaneous learning of MSA and EA (i.e., optionality in the input) caused perhaps linguistic dissonance for the IPG participants. This, in turn, hindered their comprehension of the input, resulting in error-full production (as suggested by Alhawary, 2013, 2021; Eisele, 2018; Nation & Macalister, 2010; Schwartz & Causarano, 2007). Such an argument is supported by the past negation data of the UIPG (i.e., L2ers of MSA only). That is, their simultaneous learning of the two MSA past negation constructions (*lam + imperfect verb* and *mā + perfect verb*) was a source of confusion to these participants and prevented them from making significant gains over time.

Various pieces of evidence for the confusion experienced by the IPG participants are yielded from the qualitative analysis of the errors produced in the target structures in which MSA and EA diverge. One of the errors produced in the present tense structure was due to the unauthentic nature of some production samples, in which EA-like verbs (e.g., *yarūh* “to go”)

were produced without the habitual prefix *bi-*, as was illustrated in (69a–b, Section 4.3.1.2: Chapter 4). Another piece of evidence is that some IPG participants did not supply the habitual marker *bi-* in the EA verbal present negation construction and produced the EA-like verb *yišūf* “to see” in an MSA negation construction, as was illustrated in (95a–c, Section 4.3.2.2: Chapter 4). In fact, this supports the concern raised by Alhawary (2021), who maintained that exposure to two varieties from the outset results in unauthentic speech never attested in the discourse of adult native speakers of Arabic. More importantly, some IPG participants produced hybridized negators, which have no presence at all in the negation paradigm of any Arabic varieties. This was evident in their use of the enclitic *-š* of the discontinuous EA negator *ma-š* with the MSA negator *lā* and with the EA non-discontinuous negator *miš*, as was illustrated in (96a–b, Section 4.3.2.2: Chapter 4). This finding, on the one hand, provides counterevidence to the findings reported in Shiri and Joukhadar’ (2018) study, which found the simultaneous teaching of MSA and CA (but in different sections) to be advantageous to first-year Arabic L2ers who gained knowledge about the differences between the two varieties, reached a high level of accuracy, and hardly ever used any hybrid forms that would render their use of the two varieties unnatural or unauthentic. On the other hand, the current finding confirms the argument that exposure to more than one variety from the outset results in morphosyntactically hybridized or pidginized discourse never produced by native speakers of Arabic (Alhawary, 2013).

Other evidence for the linguistic dissonance experienced by the IPG participants stem from the errors attested in their production of the complex subjunctive and cause–effect subordination structures. Recall, like English, MSA requires spelling out the complementizer *?an* “to” in the subjunctive subordination structure (+ complementizer), whereas EA entails dropping this particle (- complementizer) and spelling out the prefix *bi-* with the matrix verb in the simple

and progressive present contexts (i.e., the contexts in which the questions of the subjunctive subordination task were asked). The error that was due to lack of *ʔan* “to” suppliance decreased considerably over time in the data of the UIPG, particularly at the third-year level. In contrast, it was persistent in the data of the IPG, whose second-year and third-year participants produced approximately an equal number of this error in their production of the MSA subjunctive subordination structure (see Table 4.22: Chapter 4).

Furthermore, one of the errors observed in the cause–effect subordination data was due to producing a nominal-verbless, nominal-verbal, or verbal sentence in the subordinate clause headed by *bisabab* “because of,” which is considered the first part of the *ʔiḏāfah* construction, and which must be followed by a genitive noun phrase. This error was prevalent in the data of the IPG, compared to that of the UIPG (see Table 4.26: Chapter 4). To elaborate, the UIPG participants were introduced to only the complex structure with the conjunction *liʔanna* “because,” as well as the simple structure containing the *ʔiḏāfah* construction headed by *bisabab* “because of.” In contrast, besides this structure, their IPG counterparts were also exposed to the EA conjunction *ʕalašān/ʕašān* “because,” which perhaps added a more cognitive-load on them, which is argued here to be a source for input incomprehensibility. In other words, they most likely went through confusion in terms of how to convey the cause–effect relation through the *ʔiḏāfah* construction, through the conjunction *liʔanna* “because,” and through the conjunction *ʕalašān/ʕašān* “because.”

Moreover, the error that was due to producing a verbal sentences (i.e., VSO word order) instead of a nominal-verbless sentence or nominal sentence containing verb predicate (i.e., SVO word order) in the subordinate clause headed by *liʔanna* “because” was predominately produced by the IPG group, particularly the third-year participants. This may arguably indicate that the EA

variety interfered and played a nonfacilitative role in the acquisition of this complex structure with the conjunction *li?anna* “because,” as will be discussed further in the following section.

The linguistic dissonance resulting from input optionality is further supported by the performance of the IPG and UIPG on the past negation. The nonsignificant improvement from the first-year to second-year to third-year levels within either L2 group strongly indicate that lumping the two MSA past negation constructions (*mā + perfect verb* and *lam + imperfect verb*) and introducing them early on impeded the participants’ input comprehensibility. Nevertheless, this does not explain why the UIPG significantly outperformed the IPG at the first-year and third-year levels. The only apparent reason for this asymmetrical performance can be attributed to the fact that the IPG participants were exposed to the EA verbal past negation system (*mā + perfect verb + š*) in addition to the MSA system. It can thus be argued that the simultaneous exposure to two MSA past negation constructions was overwhelming to the UIPG participants, but the additional exposure to the EA system was cognitively more overwhelming to the IPG participants. This is evidenced by the fact that the accuracy mean scores of the IPG on this structure were less than 20% across all three proficiency levels (see Table 4.14 and Figure 4.7: Chapter 4) and that the third-year IPG participants significantly underperformed the first-year UIPG participants. The very weak accuracy performance of the IPG participants may indicate that they did not formulate any rule about the past negation in either variety and perhaps ceased to develop in their acquisition of this structure—even after 3 years of learning Arabic. Hence, such a finding can possibly signify that early fossilization might have taken place among the IPG in their acquisition of the past negation (i.e., local fossilization). Further discussions on how the optionality in the input (i.e., input obscurity) hinders facilitative transfer from the L1 are provided in the following section.



The notion of input frequency and presentation timing of the target structures in the instructional textbooks can also account for some of the findings. Recall, the target morphosyntactic structures examined in the current study are considered among the most high-frequency structures. To ensure the L2 participants were exposed to these structures from an earlier stage, the frequency and presentation timing of these structures were analyzed in the instructional textbooks used by both L2 groups across all proficiency levels (see Section 3.4.1: Chapter 3). The data seem to offer compelling evidence for the frequency and timing effects.

The strongest piece of evidence in support of the input frequency factor comes from where Section 5.2.1.4. ended. The IPG participants opted predominately for the MSA structures rather than their EA counterparts. This can be well understood considering the frequency of the MSA and EA target structures in the input they received: Part I and Part II of the third edition of *Al-Kitāb fī Ta'allum Al-'Arabiyya* (Brustad et al., 2011, 2013). To illustrate, the MSA present tense inflected for the 3.s.m and 3.s.f was formally introduced in Lesson 4 and was frequently recycled through Lesson 23 (i.e., the last lesson) of their instructional textbook (see Table 3.6: Chapter 3). In contrast, the EA present tense (i.e., *bi + imperfect verb*) was formally introduced in Lesson 4 but was not sufficiently recycled thereafter (i.e., less than four times per lesson) and was not recycled at all in some lessons, particularly in the 3.s.f context. With regard to the past and present negation systems (see Tables 3.7 and 3.8, respectively: Chapter 3), the MSA constructions (*mā + perfect verb* and *lam + imperfect verb*) were formally introduced in Lessons 8 and 13, respectively. Although the former was not sufficiently recycled afterwards, the latter was consistently recycled thereafter. The MSA verbal present negation construction (*lā + imperfect verb*) was formally introduced in Lesson 4 and was sufficiently recycled thereafter. In contrast, the EA past negation construction (i.e., *ma + perfect verb + š*) and the verbal present

negation system (i.e., *ma + bi + imperfect verb + š*) were neither formally introduced nor sufficiently recycled throughout almost all lessons of the textbook used by the IPG. As for the subjunctive subordination structure (see Table 3.9: Chapter 3), the MSA structure (+ complementizer) was formally introduced in Lesson 10 and sufficiently recycled thereafter. However, its EA counterpart (- complementizer) was formally introduced in Lessons 4 and 10 but was not recycled more than four times in almost all lessons, except for Lessons 8, 12, and 13. Regarding the cause–effect subordination (see Table 3.10: Chapter 3), the conjunction *liʔanna* “because” was formally introduced in Lesson 6 and sufficiently recycled starting from Lesson 10. In contrast, the EA conjunction *ʕašān/ʕalašān* “because” was neither formally presented nor sufficiently recycled throughout all the lessons.

Thus, the IPG participants did not opt for the EA variety in their production due to the insufficient frequency of its structures in the input (i.e., instructional textbook). The IPG’s adherence to MSA points out that they resorted to the external resource that is available to them in the input (i.e., MSA), as expected by the CDST (de Bot et al., 2007; de Pot & Larsen-Freeman, 2011; Larsen-Freeman, 1997, 2017). More importantly, the paucity of occurrence of EA structures in the input may have led these participants to avoid this variety in their production. Such avoidance may strongly indicate that learning/acquisition of EA did not take place among these participants across all proficiency levels. This assumption is borne out by researchers who subscribe to input-driven accounts of language acquisition and who argue that the development of L2 structures is fundamentally sensitive to their frequency in the input—that is, the more the L2 forms are frequently present in the input, the more likely they are acquired by L2ers (e.g., Barlow & Kemmer, 2000; Bybee, 2006; Bybee & Hopper, 2001; Ellis, 1994, 2002a, 2002, 2012; Gass & Selinker, 2008; Long, 1996; Nation & Macalister, 2010; Schmid, 2017).

Another piece of evidence that seems to predictively support the role of input timing and frequency comes from the past tense data. Given that MSA, EA, and English exhibit a past tense feature, the IPG and UIPG would be expected to perform comparably at all proficiency levels on this structure. However, the latter significantly outperformed the former at the first-year and second-year levels. In addition, the error that was due to producing present tense forms instead of their past counterparts was prevalent among the first-year and second-year participants of both L2 groups, whereas it was rarely produced by the third-year participants of both L2 groups (see Tables 4.5 and 4.6: Chapter 4). The analysis of the presentation timing and frequency of the past tense in the two instructional textbooks used by both L2 groups can perhaps account for these two findings. As shown in Table 3.5 in Chapter 3, the past tense was formally introduced (with both forms of 3.s.m and 3.s.f) in Lesson 8 of both textbooks. However, while it was sufficiently and consistently recycled thereafter in the second edition of *Al-Kitāb* (i.e., the textbook used by the UIPG), it was not introduced at all in Lesson 9 and was introduced in the 3.s.m context only in Lesson 10 of the third edition of *Al-Kitāb* (i.e., the textbook used by the IPG). Accordingly, the lack of presentation of the past tense in Lessons 9 and 10, after it was formally introduced in Lesson 8 in the IPG's textbook, might explain why the accuracy mean scores of the first-year and second-year IPG participants were below the 20% range. Possibly, the lack of exposure to this tense for at least 2–3 weeks after its formal introduction prevented these participants from formulating any rule regarding how to form and use the past tense structure. However, with more exposure, the IPG participants were eventually able to produce the past tense more properly and accurately, particularly at the third-year level.

Taken altogether, despite the presence of the tense feature in the participants' L1 (English), the first-year and second-year participants exhibited a tendency to converge somewhat

in their use of the present tense as the default form when narrating past tense events. This finding is congruent with that reported in Alhawary's (2019) study, which found a similar pattern among the nonintensive group of Chinese L1 first-year and second-year Arabic L2ers, but it does not align with that reported in Al-Hamad's (2003) study, where Russian L1 and Chinese L1 learners of Arabic as an L2 were extremely successful in their production of the past tense from early on. However, the considerable decline of the tense error at the third-year level of both L2 groups in the current study highlights the superior advantage of the input frequency effects over cross-linguistic effects, and this result is consistent with that reported in Alhawary's (2019) study, which found a similar effect of input exposure on the acquisition of past tense by Russian L1 and Chinese L1 speakers with intensive exposure to Arabic as an L2.

An important question that remains unanswered is why the participants of both L2 groups experienced a delay in their acquisition of the past tense, compared to its present counterpart. The answer can be again attributed to the presentation timing of these two structures in the input. To elaborate, although the past tense was formally introduced in Lesson 8 of both textbooks used by the IPG and UIPG (see Table 3.5: Chapter 3), the MSA present tense was introduced at an earlier stage, particularly in Lesson 4 of both textbooks, and was consistently recycled thereafter (see Table 3.6: Chapter 3). Plausibly, this variation in the presentation timing of both tense structures in the input reflects the asymmetrical accuracy performance of both L2 groups across all proficiency levels on both tense structures. That is, the sequence of presentation (along frequency) of them in the input is argued to be the main cause for the higher accuracy rate of all L2 participants on the present tense, as opposed to the past tense. The implication here is that the structure that was presented earlier (present tense) was acquired more readily and accurately than the later-introduced structure (past tense). This effect of input exposure is consistent with

previous findings reported in Alhawary's (2009a) study, where Arabic L2ers from English, Spanish, and Japanese L1 backgrounds (across all proficiency levels) were more accurate in their production of the past tense than the present tense because the former was introduced in the input much earlier than the latter. A similar effect of input exposure was likewise yielded from Alhawary's (2019) data, which revealed the following: (a) Chinese L1 speakers with nonintensive exposure to Arabic performed comparably on the past tense and present tense because both structures were introduced at the same time in their instructional textbook; (b) Chinese L1 speakers with intensive exposure to Arabic performed better on the present tense than the past tense because the former was introduced in the input before the latter; (c) Russian L1 first-year and second-year L2ers of Arabic with nonintensive exposure, as well as Russian L1 first-year L2ers of Arabic with intensive exposure, underperformed their Chinese counterparts on the present tense because it was introduced at a later stage in the former participants' textbook.

Another finding that seems to predictively provide evidence to the role of presentation timing of the target structures in the input comes from the past negation data. Recall, the participants of IPG and UIPG across all proficiency level exhibited a tendency to use the *mā + perfect verb* construction more often than the *lam + imperfect verb* construction. This finding is congruent with those reported in Albirini et al., (2019) and Alhawary's (2009a, 2013) studies, which found a similar tendency among English L1 learners of Arabic as an L2. A possible explanation for this finding is that the *mā + perfect verb* construction was prevalent in the data because of its simple nature, which entails spelling out the negator, without any change in the perfect form of the verb (Albirini et al., 2019; Alhawary, 2009a, 2013). However, a more convincing explanation can be attributed to the presentation timing of both negation constructions in the instructional textbooks used by the participants. As shown in Table 3.7 in

Chapter 3, the *mā* + *perfect verb* construction was formally presented in Lesson 8 of both textbook editions, but it was not sufficiently recycled thereafter. In contrast, the *lam* + *imperfect verb* construction was introduced at a later stage in Lesson 13 of both textbook editions and was sufficiently recycled afterwards, particularly in the third edition (i.e., the textbook used by the IPG group). It was expected that the first-year participants would exclusively use the *mā* construction because it was the only form to which they were exposed during the time of data collection. However, this does not explain why it was used predominately by the second-year and third-year participants. The argument here is that, because the *mā* construction was learned first, these participants perhaps did not make any cognitive attempts to learn the *lam* construction because they likely perceived it as an optional (and presumably unneeded) structure that conveys the same meaning as the *mā* construction, which they had learned previously. This occurred despite the fact that the *lam* construction is less marked, more frequently recycled in the textbooks, and more similar to the negation construction exhibited by their L1 (English). In addition to this sequence of presentation (i.e., the *mā* construction first then the *lam* construction), the insufficient recycling of the former is another feasible reason that might explain why the L2 participants' accuracy on the past negation did not show any significant improvements over time. They ceased to develop in their acquisition of this construction due to its insufficient recycling in the input.

Other evidence in support of the role of input frequency comes from the proficiency factor (i.e., more exposure to the L2 forms from) within the UIPG only. Recall, the UIPG improved significantly as their proficiency increased, particularly at the third-year level, in their acquisition of the present tense, present negation, subjunctive subordination, and cause–effect subordination. However, as discussed previously, if there is optionality in the input,

comprehension might be hindered. As a consequence, this might minimize the facilitative role of input frequency (the proficiency factor in this case) in the acquisition of L2 structures. Evidence for this stems from the past negation findings, which revealed that the UIPG participants did not make any significant gains in their acquisition of the MSA past negation. This is also supported by the fact the IPG participants did not improve significantly from the first-year to third-year levels in their acquisition of the target structures in which MSA and EA diverge: present tense, past negation, verbal present negation, subjunctive subordination, and cause–effect subordination. Thus, more exposure to the L2 forms would be ineffective if comprehension does not take place, as theoretically maintained by Alhawary (2013, 2021), Corder (1967), Gass and Selinker (2008), Krashen (1985), Long (1996), and Nation and Macalister (2010)—and as empirically evidenced by Loschky (1994), Sachs et al. (1981), and Snow et al. (1976). In other words, optionality or obscurity in the input (i.e., learning two distinct structures in one variety or learning two or more structures in two varieties) is most likely the reason for the statistically flat performance of both the UIPG participants on the past negation and the IPG on the target structures in which MSA and EA diverge—despite the prolonged years of exposure to Arabic.

#### ***5.2.2.2 L1 transfer***

As discussed in Chapter 2, the input effect cannot work in isolation from the L1 transfer effect, and this was well captured by Gass and Mackey (2002), who stated that “research on the role of the native language in SLA has shown that it can impact learning in both positive and negative ways. It is possible that frequency effects are intertwined with native language effects in relation to some aspects of learning. (p. 256; see also Alhawary, 2005, 2009a, 2009b, 2013, 2019; Azaz, 2016; Ellis, 2002b; Lin & Alhawary, 2018; Long & Sato, 1983). Certainly, the current data confirm this assumption. The strongest evidence in support of the role of L1 transfer comes again

from the predominant use of MSA in the data from the IPG group. As postulated by the CDST (de Bot et al., 2007; de Pot & Larsen-Freeman, 2011; Larsen-Freeman, 1997, 2017), this finding was predicted, given that the development of the L2 (Arabic) is restricted by the L1 (English), one of the resources available to L2ers (the IPG participants in this case). To elaborate, based on the generalization of the typological pairings on the acquisition of—or at best the performance on—the target morphosyntactic structures, the IPG’s strict adherence to the MSA variety was somehow expected due to the structural proximity between MSA and English, on the one hand, and the structural differences between EA and English, on the other hand.

However, despite the predominant presence of MSA in the data of the IPG, the additional and simultaneous exposure to EA (i.e., obscurity in the input) obstructed positive/facilitative transfer from English. Several findings support this argument. First, of all the target structures, the IPG participants showed a significant improvement over time in their acquisition of the past tense only, in which MSA, EA, and English converge, whereas they did not show any significant improvements, particularly from the first-year to third-year levels, in their acquisition of the other target structures, in which MSA and English behave differently from EA. Second, although the third-year UIPG participants were statistically on par with the control participants in their accuracy on the past tense and present tense (in the 3.s.m context only), as well as on the verbal present negation structure, the third-year IPG participants did not achieve native-like proficiency in any of the target structures. Third, although the error that was due to authenticity in the production of the subjunctive subordination (i.e., lack of *ʔan* “to” suppliance) decreased considerably as proficiency increased within the UIPG, it did not exhibit such a decreasing tendency as proficiency increased within the IPG (see Table 4.22: Chapter 4). In fact, only the data of the UIPG were somehow consistent with those reported in Alhawary’s (2009a) study,



which found that the subjunctive particle *ʔan* “to” emerged at an early stage in the interlanguage systems of English L1 beginner L2ers of Arabic, and this was expected due to the structural proximity between MSA and English in spelling out the subjunctive particle (+ complementizer). However, the persistence of this error in the data of IPG across all proficiency levels suggests that facilitative transfer from the participants’ L1 (English) was plausibly hindered due to their additional exposure to EA, in which the subjunctive particle is null (- complementizer). Fourth, notwithstanding the structural proximity between Arabic and English in terms of the suppliance of a nominal-verbless sentence or nominal-verbal sentence (i.e., SVO) in the subordinate clause headed by the conjunction *liʔanna* “because”, this structure posed some difficulties for the IPG participants, who exhibited a tendency to produce a verbal sentence (i.e., VSO) following *liʔanna* “because” (see Table, 4.26: Chapter 4). The most convincing reason for this error can be likely attributed to interference from EA, which exhibits the conjunction *ʕalašān/ʕašān* “because” followed by either a nominal or verbal sentence. Plausibly, the IPG participants overgeneralized and applied the rule of the conjunction *ʕalašān/ʕašān* when using the conjunction *liʔanna* “because”—signifying negative interference from the EA *ʕalašān/ʕašān* construction in the acquisition of its *liʔanna* counterpart.

The central role of input transparency in facilitating positive transfer from the L1 is further supported by the data of the UIPG as well. Recall again that, unlike other target structures, the UIPG participants did not show any significant improvements in their acquisition of the past negation. It was argued previously that optionality in the input (i.e., the *mā* construction and the *lam* construction), presentation timing of the two constructions, and lack of input frequency of the former construction were most likely the causes for the flat performance of these participants across all proficiency levels. However, based on the generalization of the

typological pairing of the acquisition of the past negation, the *lam* construction was expected to be acquired more readily due to its structural proximity in MSA and English. That is, the two negators in MSA and English (*lam* and *did not*, respectively) head a negation phrase higher than the verbal phrase being negated, and the verb is produced in its imperfect form. However, the UIPG participants (and the IPG) did not opt for this construction in their production, as discussed in the preceding section, and when they attempted to do so, they mostly produced the negator *lam* with perfect verbs (see Table 4.15: Chapter 4). Thus, exposure to two structures in one variety (as revealed by the acquisition data from the UIPG on the past negation) and exposure to two varieties (as revealed by the acquisition data from the IPG on the structures in which MSA and EA diverge) seem to prevent facilitative transfer from the L1 and hence are detrimental to the acquisitional development of L2ers. This argument is borne out by various L1 transfer hypotheses, including the FRH (Hwang & Lardiere, 2013; Lardiere, 2008, 2009), FT/FAH (Schwartz & Sprouse, 1996), and robustness of L1 transfer hypothesis (RL1TH: Alhawary, 2021), which predict that input transparency is a prerequisite for successful (or positive) L1 transfer to take place in the acquisition of an L2.

Cross-linguistic effects from the L1 in relation to the notion of UG access are further supported by the past and present tense data, particularly the subject-verb (S-V) agreement data. Recall, the third-year UIPG participants and the control participants performed comparably on the past tense and present tense in the 3.s.m context but differed significantly in the 3.s.f context. In addition, the qualitative analysis of the errors produced by the IPG and UIPG in the past tense structure revealed that, while the agreement error was the least in the 3.s.m context (see Table 4.5: Chapter 4), it was the predominant error produced in the 3.s.f context and did not exhibit a decreasing tendency from the second-year to third-year levels within either the IPG or UIPG (see

Table 4.6: Chapter 4). Similarly, both L2 groups were highly accurate in inflecting the present tense for the proper agreement form in the 3.s.m context (see Table 4.10: Chapter 4), but they exhibited a tendency to converge in their use of the default form (3.s.m) in their production of this tense in the 3.s.f context (see Table 4.11: Chapter 4). However, although this agreement error decreased as proficiency increased within the UIPG, it was prevalent in the data of the IPG participants and did not show a decreasing trend over time. These findings were predicted based on the typological pairings, postulating that the acquisition of the S–V agreement feature would be problematic for the L2 participants because of its absence or noninstantiation in their L1 (English). This confirms previous findings that found the acquisition of this feature to be problematic for Arabic L2ers whose L1s were English, French, and Japanese (Alhawary, 2009a); English and French (Alhawary, 2005, 2009a, 2009b); and Russian and Chinese (Alhawary, 2019). These participants exhibited higher accuracy on the 3.s.m than the 3.s.f agreement forms and used the former, as the default form, in the context of the latter.

The current findings have certain implications for various L1 transfer hypotheses that attempt to account for ultimate attainment in the acquisition of inflectional features in light of UG access, particularly the FFFH (Hawkins, 2001; Hawkins & Chan, 1997), the interpretability hypothesis (Hawkins, 2005; Hawkins et al., 2008), FRH (Hwang & Lardiere, 2013; Lardiere, 2008, 2009), FT/FAH (Schwartz & Sprouse, 1996), and RL1TH (Alhawary, 2021). According to the FFFH (Hawkins, 2001; Hawkins & Chan, 1997) and interpretability hypothesis (Hawkins, 2005; Hawkins et al., 2008), L2 uninterpretable  $\phi$ -features (the Arabic S-V agreement feature in this case) that are not instantiated in the L1 (English in this case) are predicted to be permanently impaired, so their acquisition is impossible after the critical period, regardless of the complex or simple nature of the target features, and irrespective of the length of exposure to the L2 input.

Based on this prediction, it would be supposed that none of the L2 participants of both L2 groups (IPG and UIPG) across all proficiency levels would perform at or near the ceiling level (100% or within 80% of accuracy, respectively) on the past tense and present tense in both the 3.s.m and 3.s.f agreement contexts. However, upon examining the performance of every participant, the prediction of permanent impairment did not hold. To illustrate, for the past tense inflected for the 3.s.f agreement form, the accuracy of five participants (two in the IPG and three in the UIPG) was at the 100% level, and the accuracy of seven participants (one in the IPG and six in the UIPG) was within the range of the 80–95% level. Similarly, for the present tense inflected for the 3.s.f agreement form, the accuracy of two participants (one in the IPG and one in the UIPG) was at the 100% level, and the accuracy of 12 participants (four in the IPG and eight in the UIPG) was within the range of the 80–95% level. These findings align well with previous research that found a similar pattern of high accuracy rate on the S-V agreement among Arabic L2ers from various L1 backgrounds that do not exhibit a distinction in verbal gender agreement, including English, French, and Japanese (Alhawary, 2009a); English and French (Alhawary, 2005, 2009a, 2009b); Russian (in the present tense) and Chinese (Al-Hamad, 2003; Alhawary, 2019).

As for the FRH (Hwang & Lardiere, 2013; Lardiere, 2008, 2009), it predicts that the acquisition of the uninterpretable  $\phi$ -feature (the Arabic S-V agreement feature in this case) that is not instantiated in the L1 (English in this case) is eventually possible. However, this hypothesis predicts that the acquisition of this feature is expected to involve a complex process that is contingent upon the L2ers' ability to reconfigure/reassemble it from the L1 onto the L2. The current data confirm the first prediction of this hypothesis because performance at the ceiling level was eventually attained by some L2 participants, as discussed in the preceding paragraph. However, the second prediction of the FRH was not confirmed by the data because the learning

of Arabic S-V was expected to be simple rather than complex (at least on the surface). The supposed simplicity is mainly due to the phonological clue (or regularity) that is realized as a zero morpheme for the 3.s.m and as the suffix *-at* for the 3.s.f on the past tense verb (e.g., *kataba ʔaṭ-ṭālib* “the male student wrote” versus *katab-at ʔaṭ-ṭālib-a* “the **female** student wrote”) and that is realized as the prefix *ya/ya-* and the prefix *ta/ta-* for the 3.s.m and the 3.s.f, respectively, on the present tense verb (e.g., *ya-ktub ʔaṭ-ṭālib* “the **male** student writes” versus *ta-ktub ʔaṭ-ṭālib-a* “the **female** student writes”). Accordingly, because of the S-V agreement simple regularity or the rhyming effects on past and present verbs, “it would be unwarranted to assume complexity of feature realization (which does not exist in this case) to result in added learning burden or difficulty” (Alhawary, 2019, p. 152). Hence, this would leave room to account for the difficulty in the production of the past tense and present tense in light of the FT/FAH (Schwartz & Sprouse, 1996), and more precisely the RL1TH (Alhawary, 2021), which argue for the critical role of *presence* or *absence* of a feature in the development of the L2ers’ interlanguage systems. In other words, notwithstanding the simplicity of the S-V agreement feature in Arabic (cf. the FRH), its absence in English was the main cause for the difficulty encountered by the current L2 participants (à la the FT/FAH and RL1TH).

However, this is not to say the L2 participants were not aware of the Arabic S-V gender agreement. Given that the tense feature is conflated with the agreement feature in Arabic (e.g., the suffix for the tense feature is also considered as an agreement feature on the past tense), the presence of the tense feature in English raised the participants awareness to produce the S-V agreement feature (see also Alhawary 2009a, 2019). This argument is supported by the qualitative analysis of the agreement error produced, where the participants made some attempts to produce the 3.s.f agreement feature by conjugating the verbs (particularly in the past tense

context) for the 1.s, 2.s.m, and 2.s.f forms instead of the 3.s.f form. However, those attempts were often unsuccessful due to the fact that English exhibits an impoverished agreement paradigm—that is, it lacks a verbal gender agreement feature—which highlights again the critical role played by the presence and absence of a feature in the L1 in the acquisition of an L2. Equally important, the absence of diacritics in the input can explain why the L1 transfer effects could not be minimized, particularly in the acquisition of the 3.s.f agreement feature in the past tense context, as will be discussed immediately below.

### ***5.2.2.3 Implicit and explicit learning***

It is well-established that deliberate focus-on-form instruction (Azaz, 2016; Norris & Ortega, 2001; Robison, 1996; Spada & Tomita, 2010) and visual presentation L2 forms in the input (Al Midhwah & Alhawary, 2020; Lin & Alhawary, 2018; Kim & Godfroid, 2019) are considered among the explicit learning techniques that make the input salient to L2ers and draw their attention to the gaps in their interlanguage systems in comparison to the target L2. According to the noticing hypothesis, conversion of *input* into *intake* is contingent upon noticing it consciously (Schmidt, 1990). By extending this notion to the implication of the implicit–explicit continuum for the current study, it is assumed here that the incorporation of focus-on-form instruction and visual presentation of the target structures in the input (i.e., explicit learning techniques) would be essential for developing robust implicit knowledge, characterized at least as the accurate use of Arabic by the L2 participants. In contrast, leaving the participants with focus-on-function instruction and aural presentation of the target structures in the input (i.e., implicit learning techniques) would lead us to predict that it would not yield a similar effect of developing implicit knowledge. This assumption was somehow confirmed by the current findings.

Recall again that the IPG participants opted predominately for MSA rather than EA in their description of the target pictures. This was justified previously due to the frequent and sufficient recycling of the MSA target structures in the input (i.e., instructional textbook) and due to the proximity between MSA and English in the target structures. However, and in addition to these two factors, the lack of explicit focus-on-form instruction of the EA structures is argued here to be another critical factor behind this finding. To elaborate, of all the target structures in which MSA and EA diverge, only the EA present tense (see Table 3.6: Chapter 3) and EA subjunctive subordination (see Table 3.9: Chapter 3) were explicitly introduced in the textbook used by the IPG, whereas the EA past negation (see Table 3.7: Chapter 3), EA verbal present negation (see Table 3.8: Chapter 3), and EA cause–effect subordination (see Table 3.10: Chapter 3) were implicitly introduced. This finding confirms previous findings that found explicit instruction to be more effective and beneficial than its implicit counterpart in the development of L2 forms (Norris & Ortega, 2001; Robison, 1996; Spada & Tomita, 2010)—and more importantly in the development of forms that are absent or assembled/configured asymmetrically in the L1 and L2 (Azaz, 2016). Additionally, this finding supports the integration of the form and function principle, which suggests that focus on grammar is as essential as focus on other skills (Alhawary, 2013). This principle stems from the communicative approach itself, which strongly accentuates the importance of focus on both function and form without marginalizing either one in favor of the other (Thornbury, 1999). However, such a principle seems to be neglected by some Arabic L2 pedagogues, especially those who call for the simultaneous teaching and learning of MSA and CA (i.e., the IA), without providing any methods and techniques that promote learning/acquisition of the two varieties from early on (e.g., Al-Batal, 1992, 1995, 2018; Al-Batal & Glakes, 2018; Belanp, 2006; Huntley, 2018; Husseinali, 2006; Isleem, 2018; Kuntz

& Benlap, 2001; Leddy-Cecere, 2018; Nassif, 2018; Palmer, 2007, 2008; Shiri, 2013, Shiri & Joukhadar, 2018; Younes, 1990, 2015).

In fact, explicit instruction seems to be critical to the development of some of the EA target structures, especially those involving very complex rules to which the L2ers' attention need to be directed deliberately. For example, the EA past negation entails attaching the discontinuous particle *ma-š* as a proclitic and enclitic on perfect verbs (e.g., *ma-zakir-š* "he **did not** study"). More complexly, the EA verbal present negation necessitates attaching the discontinuous particle *ma-š* on imperfect verbs that must be preceded by the habitual marker *bi-* (e.g., *ma-bi-yi-zākir-š* "he **does not** study"). Syntactically, these two negation structures entail a very complex operation that involves moving the verb to a negation phrase higher than the verbal phrase before merging/scrambling the verb with the discontinuous negator to convey the negation (Albirini & Benmamoun, 2015; Aoun et al., 2010).

Explicit instruction of the two EA negation structures is assumed to be needed even more in making the verbal gender agreement more salient to the L2ers. Recall, the 3.s.f agreement feature (along with the tense feature), for example, is realized in Arabic as a suffix on perfect verbs for the past tense (e.g., *zakar-it* "**she** studied") and as a prefix on imperfect verbs for the present tense (e.g., *bi-t-zākir* "**she** studies"). When negating these two tense structures, the prefix and suffix of the 3.s.f are not on the opposite ends of the verbs any more. Rather, they are blended in the middle, which makes them perceptually nonsalient (e.g., *ma-zakar-it-š* "**she** did not study," *ma-bi-t-zākir-š* "**she** does not study"). Thus, the complex rule of the EA negation, especially in the absence of explicit instruction, is plausibly another factor that can explain why the IPG participants produced only two completely correct tokens for the past negation and two completely correct tokens for the verbal present negation in the 3.s.m context only. Moreover,



the absence of explicit instruction of the syntactic differences between the two conjunctions *ʕalašān/ʕašān* and *liʔanna* “because” perhaps led the IPG participants to generalize and extend the rule of the former (i.e., *ʕašān* + a nominal-verbless, nominal-verbal, or verbal sentence) to the latter (i.e., *liʔanna* + a nominal-verbless or nominal-verbal sentence only).

However, explicit focus-on-form instruction *per se* is not sufficient if it is not accompanied by visual presentation of the forms in the input. This argument is supported by the fact that, despite the deliberate (explicit) instruction of the EA present tense and subjunctive subordination structures, the insufficient visual recycling of these two structures, along with other EA target structures, in the input did not make them perceptually salient to the IPG participants. Thus, they did not attend attentionally to these structures and did not process them easily, which in turn made them avoid the EA verity altogether in their production.

The lack of visual input of the EA structures in the input is justified due to the fact that reading and writing are almost always carried out in MSA only. As a result, the IPG participants were left with aural EA input only, which seems here to be ineffective to them. This inference is supported by previous findings that found a minimal effect of aural input (i.e., implicit technique) in the development of implicit knowledge (as in Al Midhwah & Alhawary, 2020; Kim & Godfroid, 2019; Lin & Alhawary, 2018). In addition, this inference seems to confirm another argument against the IA, particularly the integration of language skills principle. It suggests that language is best taught in an integrated manner to ensure reinforcement of all language skills, ample opportunities of input recycling, more practices on function and form, and sustainability of the L2ers’ motivation. Integration here should be across skills, not across varieties, unless sufficient time and input are provided for both types of integration at the same time, rather than short shifting one or the other (Alhawary, 2013).

Another finding that highlights the important role of visual presentation of the L2 forms in the input comes from the past tense data. As briefly pointed out above, the agreement error in production of the past tense in the 3.s.f context was the predominant error pattern due to the fact that some L2 participants of both L2 groups (IPG and UIPG) inflected the past tense for the 1.s form *-tu*, 2.s.m form *-ta*, and 2.s.f form *-ti* instead of the 3.s.f form *-at*. It is obvious that the participants were attentively aware of the agreement feature, as evident in their production of the *-t* agreement suffix. This error did not decrease as proficiency increased because the second-year and third-year participants of each L2 group produced a comparable number of this error. It is likely that the participants of both L2 groups did not receive explicit instruction focusing on the importance of the diacritics (i.e., short vowels) preceding or following the *-t* suffix and how alternation of one by another can change the agreement form—and hence changing the whole intended meaning.

More importantly, the diacritics on most of the words presented in their instructional textbooks were omitted. That is, the diacritics were mostly provided in the vocabulary list presented at the beginning of each lesson, but they were completely neglected in subsequent drills and practices. Thus, the absence of visual presentation of diacritics plausibly resulted in inaccurate production of the 3.s.f form on the perfect verbs. This result confirms previous research that found: (a) visual presentation is more effective than its aural counterpart in the development of implicit knowledge of SVO word order of an artificial language among English L1 speakers (Kim & Godfroid, 2019); (b) absence of visual presentation of stress is one of the main causes for the poor performance on Arabic stress by English L1 and Chinese L1 learners of Arabic as an L2 (Lin & Alhawary, 2018); (c) visual presentation of diacritics is more advantageous and beneficial than its aural counterpart in the development of reading fluency,

accuracy, and comprehension among English L1 beginner, intermediate, and advanced L2ers of Arabic (Al-Midhwah & Alhawary, 2020).

Taken altogether, the findings discussed in this section predictively support the notion that deliberate instruction and visual presentation of the L2 structures in the input (i.e., explicit learning enhancers) are probably among the few optimal resources available to L2ers to develop relatively robust L2 knowledge, especially in the time-limited situations allocated for them to learn an L2 (e.g., classroom settings).

### **5.3 Conclusion and Pedagogical Implications**

The current study provides compelling evidence that the IA (i.e., simultaneous learning and teaching of MSA and CA from early on) is not as effective and beneficial to Arabic L2ers. This conclusion was empirically reached by comparing the fluency, accuracy, complexity, and overall proficiency of English L1 speakers learning MSA and EA simultaneously from early on (i.e., IPG) with those of English L1 speakers learning MSA only (i.e., UIPG). Despite the great homogeneity between the two groups in terms of their L1, prior exposure to Arabic, and proficiency-placement levels, the overall findings revealed a significant advantage for the UIPG over the IPG. In addition, the IPG participants did not make any significant gains over time, particularly from the first-year to the third-year levels, in their fluency, accuracy, complexity, and overall proficiency. In contrast, their UIPG counterparts exhibited a significant improvement in the long run, particularly at the third-year level. The findings also confirmed the predictions that the simultaneous learning of two or more structures in one or two varieties, without sufficient recycling and visual (explicit) presentation of such structures in the input, can be a source of confusion, obstruct comprehension, and hinder facilitative transfer from the L1.

The current findings have some practical implications that can be beneficial to Arabic L2 teachers, textbook writers, and curriculum designers. Having a good grasp of the general acquisition tendencies among Arabic L2ers can assist in achieving a number of objectives, the most important of which is when and how to introduce the Arabic L2 forms in the input. Based on the current findings, it was observed that the simultaneous presentation and introduction of syntactically distinct and optional forms that convey a similar meaning, without sufficient recycling, explicit instruction, or visual presentation, can cause confusion for L2ers. As a consequence, they may end up with no mastery of either form. This was evident in the performance of the IPG on the target structures in which MSA and EA diverge, where the participants of this particular L2 group did not improve significantly from the first-year to the third-year levels, unlike their UIPG counterparts, who improved significantly over time in their fluency, accuracy (except the past tense with the 3.s.f form and past negation), complexity, and overall proficiency.

If the simultaneous teaching of two varieties is opted for by Arabic L2 programs, it is critically needed to frequently, explicitly, and visually introduce the structures of both varieties alike in the input (i.e., instructional textbook). However, this is most likely undoable and impractical, given the very limited time allocated for learning Arabic as an L2 in most U.S. institutions (2–3 years). To this end, the optimal solution is to exclusively adhere to one variety only for the first 1–2 years, during which the L2ers can develop “basic competence” (Abboud, 1968) or “working knowledge” (Alhawary, 2013, 2021) that can lay the groundwork for further learning of other Arabic varieties. Such an instructional practice can plausibly prevent confusion by confirming the L2ers’ learning hypothesis that they deal with one consistent set of features, forms, and structures in the L2. In turn, this would help the L2ers formulate rules about the target

language, improve substantially as their proficiency increases, and preserve their learning motivation throughout the learning journey.

Furthermore, when a variety exhibits two structures that convey one meaning (e.g., the past negation in MSA), one structure should be introduced only and sufficiently recycled afterwards in the input until it reaches a certain level of accurate automatization among the L2ers before the other structure can be subsequently introduced. Based on the notable effects found for the presentation timing of two structures in the input, where the first-introduced structure was acquired more accurately and readily than the later-introduced structure (as in the data of the past tense and present tense; see also Albirini et al., 2019; Alhawary, 2005, 2009a, 2009b, 2013, 2019), the high-frequency structure that is more similar to its counterpart in the L1 (e.g., the *lam* + *imperfect verb* construction) should be introduced before the low-frequency structure that differs morpho-syntactically from that exhibited by the L1 (e.g., the *mā* + *perfect verb* construction). This presentation can facilitate successful L1 transfer and promote accurate production of the L2 forms from an early stage of the learning process. Conversely, the opposite presentation of these two past negation structures, for example, might obstruct facilitative L1 transfer, as evident in the production of the negator *lam* “did not” with the perfect form of verbs by the L2 participants of both the IPG and UIPG across all proficiency levels.

Moreover, because it was evident that certain forms did not seem to be intensely problematic and complex to the L2 participants, particularly those exhibited in their L1 (e.g., the tense feature, the negator *lā* “does/do not”), such forms can be either explicitly or implicitly introduced from early on in the input. In contrast, other forms (e.g., the S-V gender agreement feature, the *mā* + *perfect verb* construction) seemed to pose some particular difficulties for the participants due to their absence in their L1. Thus, those forms need even more attention, and

this can be done through controlled practices and frequent recycling of these forms across all language skills in the input. Explicit instruction, preferably deductive or rule-driven instruction, is also vital. It speeds up the learning process, directs the L2ers' attention to the form more deliberately, preserves class time for practice, and "confirms many students' expectations about classroom learning, particularly for those learners who have an analytical learning style" (Thornbury, 1999, p. 30).

Visual presentation of L2 forms is equally essential, as evidenced by the production data of the past tense in the context 3.s.f, where the L2 participants attempted to produce this agreement feature by attaching the suffix *-t*. However, their attempts were unsuccessful due to their production of the 1.s feature *-tu*, 2.s.m feature *-ta*, and 2.s.f feature *-ti* instead of the 3.s.f feature *-at*. The persistence of this error was most likely due to the absence of visual presentation of diacritics (i.e., explicit learning enhancer of noticing) in their textbooks. Thus, it is important to present the L2 target forms with diacritics consistently and frequently from early on and throughout the textbook. This type of presentation can be beneficial rather than overwhelming to Arabic L2ers in developing their fluency and accuracy, as empirically evidenced by Al Midhwah and Alhawary (2021). Visual presentation of the target L2 forms in the input is certainly important when teaching any CA varieties, all of which are rarely used in formal written mediums. Exclusive reliance on aural presentation of CA is not of great benefit to the L2ers, and this might lead them to avoid the variety altogether, as revealed by the IPG data. Therefore, when teaching CA, it is practically imperative (for pedagogical purposes) to visually and frequently introduce the CA forms in the input. Taken altogether, the pedagogical techniques discussed here can be very beneficial because they can make the L2 forms salient to the L2ers and raise their noticing about the gaps in their interlanguage systems in comparison to the target

language, as these noticing mechanisms are considered one of the main prerequisites for acquiring/learning the L2 forms more readily and quickly (Schmidt, 1990).

#### **5.4 Limitations**

This study has a very few limitations that should be taken into consideration in replicated and related future research. One limitation is that several L2 participants had prior exposure to Arabic for a period up to 6 months before joining their current programs, and two of them were of Islamic heritage. Their inclusion in the study was justified due to the somewhat low number of L2ers who showed a willingness to take part in the study. However, future research should control for this variable more tightly by recruiting only those who did not have any prior exposure to Arabic. Relevantly, since only a relatively small number of participants took part in the study, future research should address this limitation by increasing the sample size of participants constituting each proficiency level within each L2 group. Nevertheless, it might not be practically possible to achieve this, especially when controlling for the various relevant variables that were controlled for in the current study, such as the participants' L1 and their prior exposure to Arabic.

Another limitation is that many of the L2 participants were unable to recall some of the key words that were indispensable for producing the target structures. However, to maintain consistency across all participants and to preserve the data from any bias or contamination, the researcher did not intervene and provided none of those words to the participants, and this led them to skip some of the target pictures and produce instead the English equivalent words. Such a limitation could have been avoided by implementing a warm-up activity to introduce or rehearse the key vocabulary before the actual tasks were carried out.

A third limitation is that the data were supposed to be collected at the end of the 2020–2021 academic year. However, because of the COVID-19 pandemic, which transformed the class-mode instruction to be virtual around the globe, the data collection was postponed to the following academic year to make the recruitment process easier. As a result, the second-year and third-year participants' data were gathered at the end of the first semester of the 2021–2022 academic year, whereas the first-year participants' data were collected at the end of the academic year to ensure that the first-year participants of both L2 groups had enough time to learn and process all the target structures. The data collection timing might have contributed to the flat performance between the first-year and second-year participants of both L2 groups on most of the target structures. This is so because the former were exposed to Arabic for a full school year, whereas the latter were exposed to Arabic for one a half years only. This limitation could have been avoided by eliciting data from all proficiency groups at the end of the school year.

For future research, it would be useful to replicate this study on other and larger pools of participants with different L1 backgrounds and on Arabic L2ers of MSA and different CA varieties other than EA. In addition, each of the SLA contributing factors examined in the current study should be searched exhaustively when investigating this topic. For example, the input frequency effect was examined in the current study by counting manually how many times each of the target structures was visually repeated in the L2 participants' textbooks. Examining the role of input frequency based on the visual presentation of the target structures may not meticulously reflect how frequently the L2 participants were exposed to these structures. To this end, future research should estimate the frequency of language use by implementing a corpus-driven analysis of the aural and visual frequency token and type of all structures targeted.



## Appendices

### Appendix A L2 Participant Language History Questionnaire

Dear participant,

You are invited to fill out the following short questionnaire. It is the first component of a research study on the learning outcomes among English L1 learners of Arabic as an L2 at different stages of their L2 development. The purpose of this questionnaire is to gather preliminary information about your language background. Any information you share will be confidential and will be used for research purposes only.

Gender:

1. Male
2. Female

Age (in years):

What is your current or most recent educational level, even if you have not finished the degree?

1. Graduate school (PhD)
2. Graduate school (MA)
3. College (BA)
4. High school
5. Other

What is your current level in the Arabic program?

1. First-year student
2. Second-year student
3. Third-year student

Is English your first/native language? If no, what is your first/native language?

1. Yes
2. No

Are you a heritage speaker of Arabic (i.e., your father or mother of an Arabic ancestor?)

1. Yes
2. No

Do you know any foreign languages besides Arabic? If yes, what are they?

1. Yes
2. No

Had you ever learned Arabic before joining your current language program?

1. Yes
2. No

If yes, how long had you learned Arabic before joining your current program?

1. 1-3 months
2. 4-6 months
3. More than 6 months

Have you lived in any Arabic-speaking countries for three months or more? If yes, which country?

1. Yes
2. No

Which Arabic variety/ies are you learning?

1. Modern Standard Arabic (MSA) only
2. MSA and Egyptian Arabic (EA)
3. MSA and Levantine Arabic (LA)

What is/are the Arabic variety/ies used by your teachers for instruction in class?

1. MSA only
2. MSA and EA
3. MSA and LA

What textbook do you use to learn Arabic?

1. Al-Kitaab fi Ta'allum Al-'Arabiyya: Second Edition
2. Al-Kitaab fi Ta'allum Al-'Arabiyya: Third Edition
3. Other. Which one?

Do you study on your own the sections designed for Al-'aammiyyah "Arabic dialects"?

1. Yes
2. No

If you study on your own the sections designed for Al-‘aammiyyah “Arabic dialects”, how many hours per week?

1. 1–3 hours weekly
2. 4–6 hours weekly
3. More than 6 hours weekly

Do you get exposed to any Arabic dialect outside the classroom (e.g., communication with friends, watching TV)?

1. Yes
2. No

If yes, how many hours per week do you get exposed to Arabic dialects?

1. 1–3 hours weekly
2. 4–6 hours weekly
3. More than 6 hours weekly

Rate your current proficiency in Arabic overall:

1. Very weak
2. Weak
3. Limited
4. Average
5. Good
6. Very good
7. Excellent

How often do you mix words or sentences from various Arabic varieties when speaking with your friends inside or outside the classroom? (This includes, for example, starting a sentence in MSA but using a word or phrase from EA or LA in the middle of the sentence)

1. Never
2. Rarely
3. Sometimes
4. Regularly
5. Often
6. Usually
7. Always

Do you understand the following sentence? If yes, please translate it into English without using any dictionary: البنت ما بتاكلش كويس زي كل يوم عشان عيانه النهارده

1. Yes
2. No

## Appendix B Control Participant Language History Questionnaire

Dear participant,

You are invited to fill out the following short questionnaire. It is the first component of a research study on the learning outcomes among English L1 learners of Arabic as an L2 at different stages of their L2 development. The purpose of this questionnaire is to gather preliminary information about your language background. Any information you share will be confidential and will be used for research purposes only.

Gender:

1. Male
2. Female

Age (in years):

What is your current or most recent educational level, even if you have not finished the degree?

1. Postdoctoral study
2. Graduate school (PhD)
3. Graduate school (MA)
4. College (BA)
5. High school
6. Other

What is your field of study/specialization?

Is English your first/native language?

1. Yes
2. No

Is Egyptian Arabic your native language?

1. Yes
2. No

How Long have you been in the United States of America?

How often do you mix words or sentences from various Arabic varieties when speaking with your friends or family? (This includes, for example, starting a sentence in Modern Standard Arabic but using a word or phrase from Egyptian Arabic in the middle of the sentence)

1. Never
2. Rarely
3. Sometimes
4. Regularly
5. Often
6. Usually
7. Always

Do you understand the following sentence? If yes, please translate it into English without using any dictionary: البنت ما بتاكلش كويس زي كل يوم عشان عيانة النهارده

1. Yes
2. No

**Appendix C** The Narrative Task of the Past Tense and Its Negation

**Figure 1c** The Past Tense and Its Negation in the 3.s.m Context: The Vacation Went from the 5th to the 17th of Last Month



الاثنين	الثلاثاء	الأربعاء	الخميس	الجمعة	السبت	الأحد
Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
		1	2	3	4	5
20	21	22	23	24	25	26
27	28	29	30	1		
				اليوم Today		

(as in Alhawary, 2009a, 2019)

**Figure 2c** The Past Tense and Its Negation in the 3.s.f Context: The Vacation Went from the 10th to the 22nd of Last Month

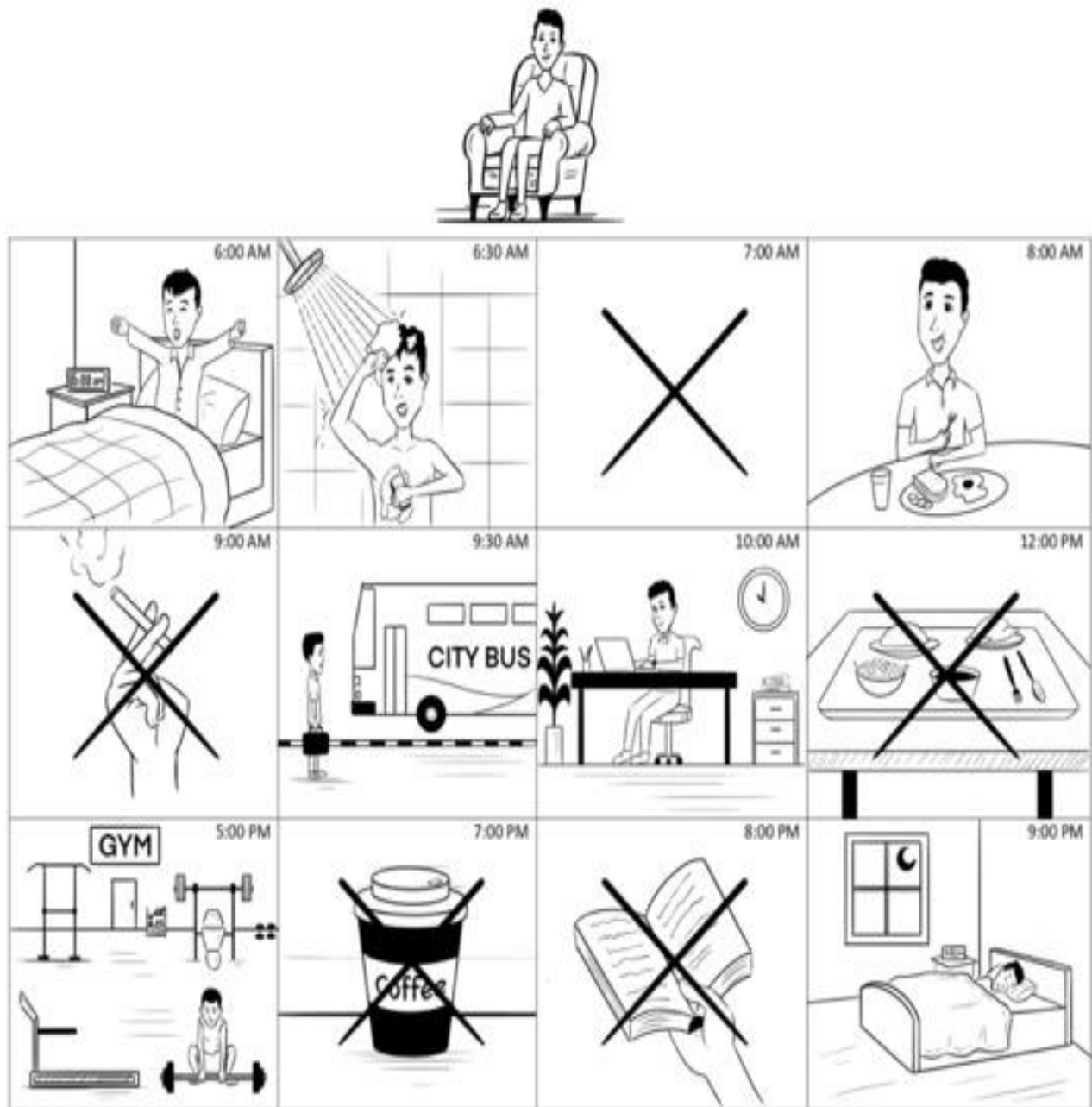


الإثنين	الثلاثاء	الأربعاء	الخميس	الجمعة	السبت	الأحد
Monday	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	1		
				اليوم Today		

(as in Alhawary, 2009a, 2019)

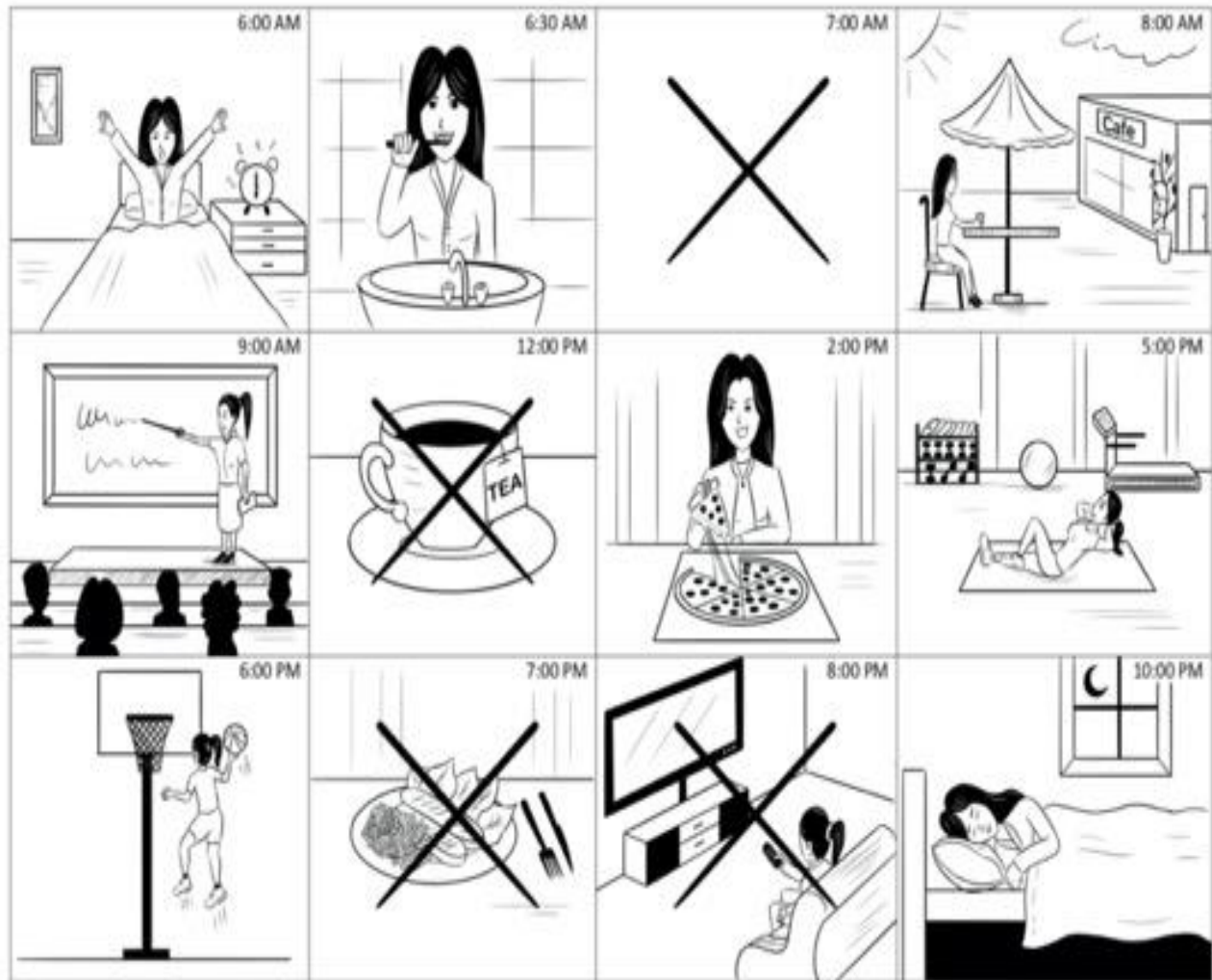
**Appendix D** The Narrative Task of the Present Tense and Its Negation

*Figure 1d* The Present Tense and Its Negation in the 3.s.m Context: The Daily Routine of a Male Character





**Figure 2d** The Present Tense and Its Negation in the 3.s.f Context: The Daily Routine of a Female Character



**Appendix E** The Manipulation Task of the Subjunctive Subordination and Cause–Effect Subordination

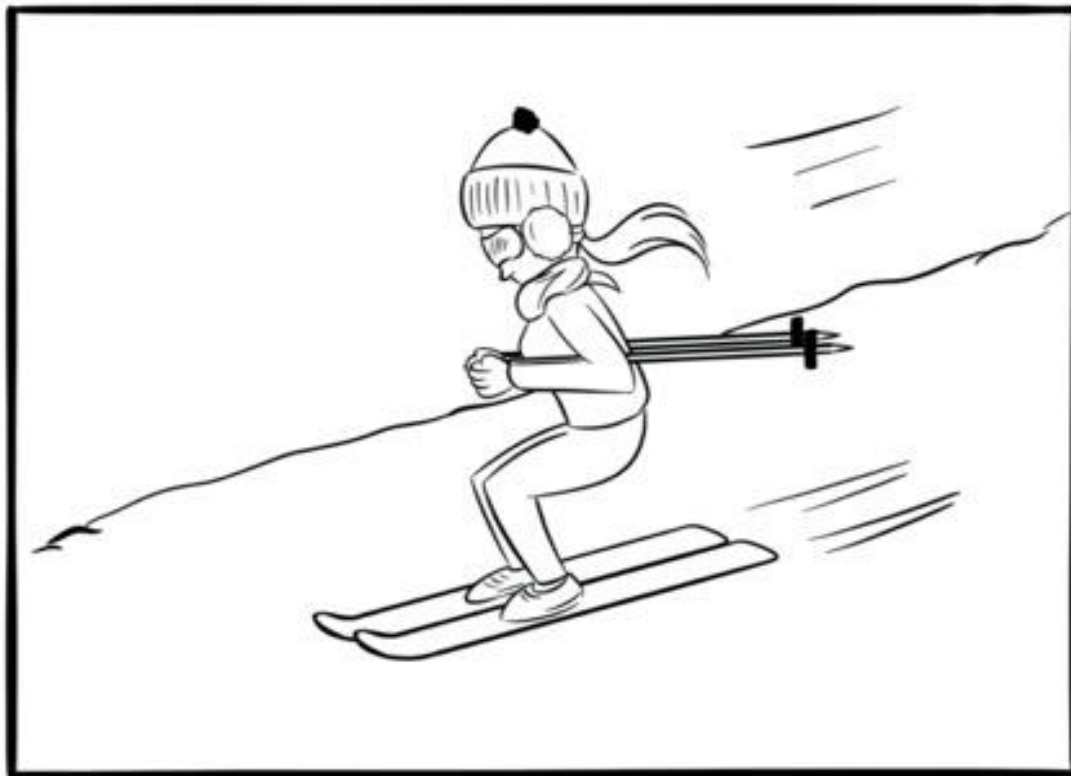
**Figure 1e** Subjunctive Subordination: What Can the Girl Do/What Is She Able to Do in Each Picture?



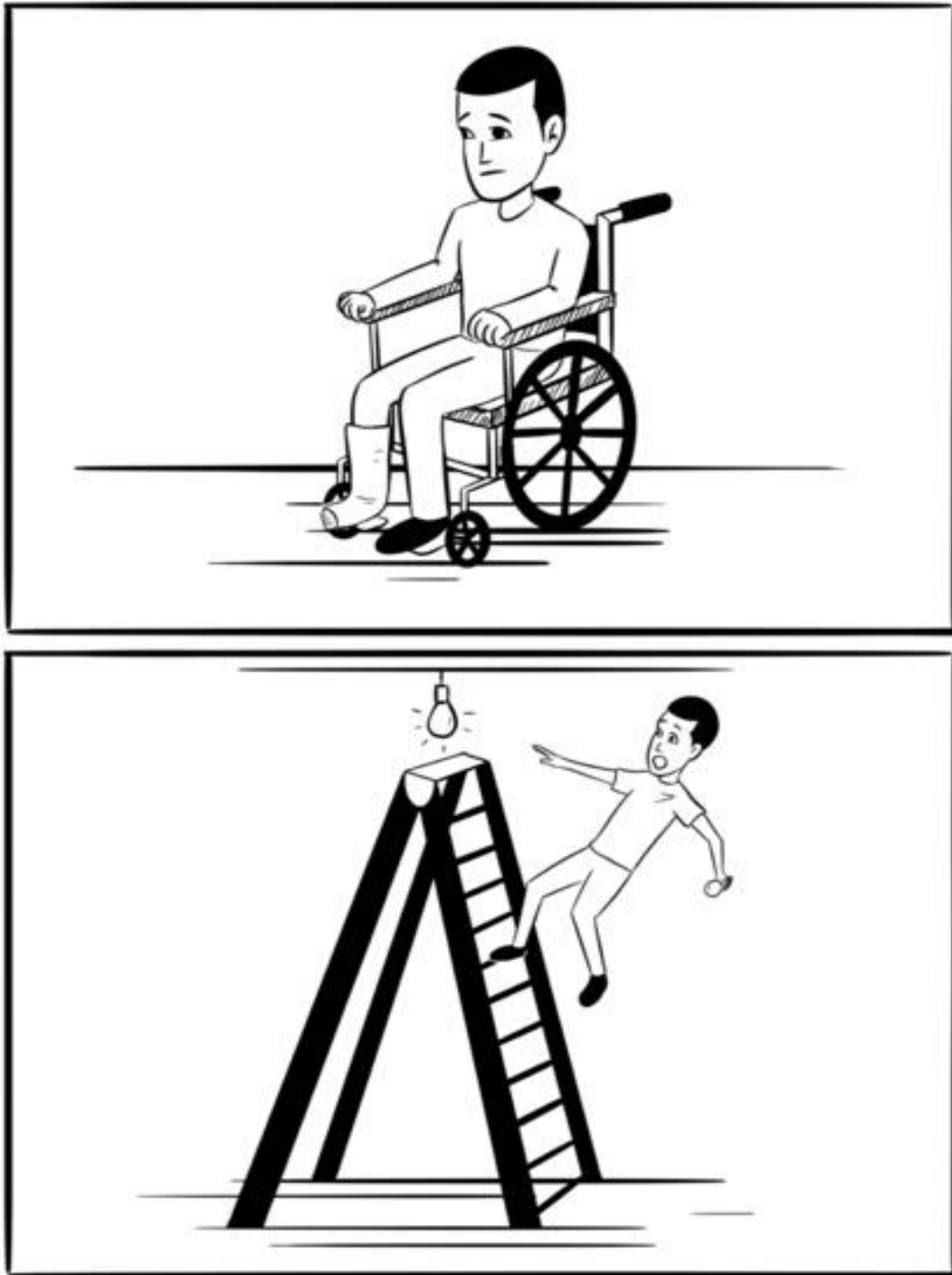
**Figure 2e** Cause–Effect Subordination: Describe What the Man Did in the First Picture and Explain Why He Did That Based on the Second Picture



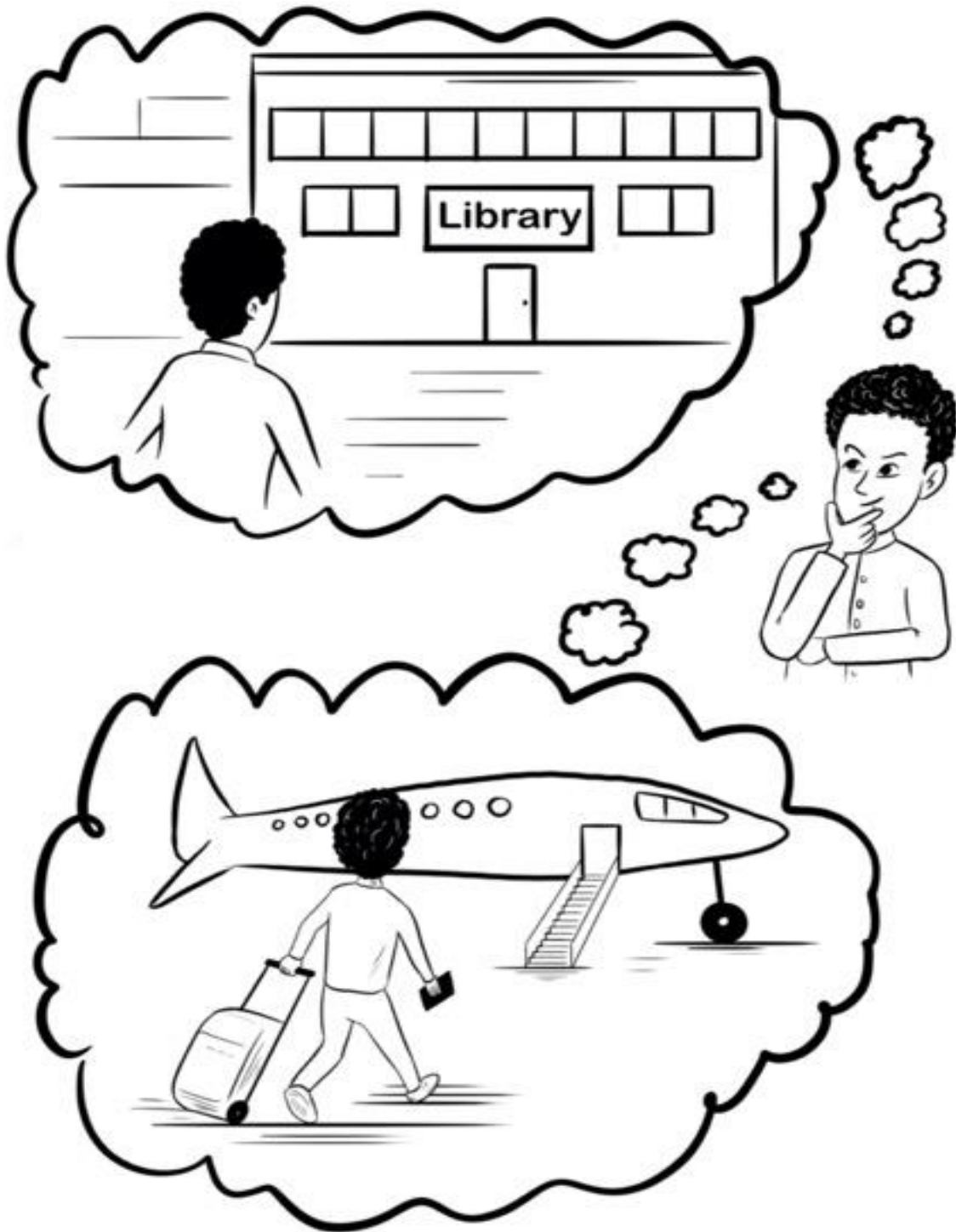
*Figure 3e Manipulation Task Distractor 1: What Is the Girl Doing in Each Picture?*



**Figure 4e** Cause–Effect Subordination: Describe the Situation of the Man in the First Picture and Explain Why He Is Like That Based on the Second Picture



*Figure 5e Subjunctive Subordination: What Is He Thinking to Do in Each Picture?*

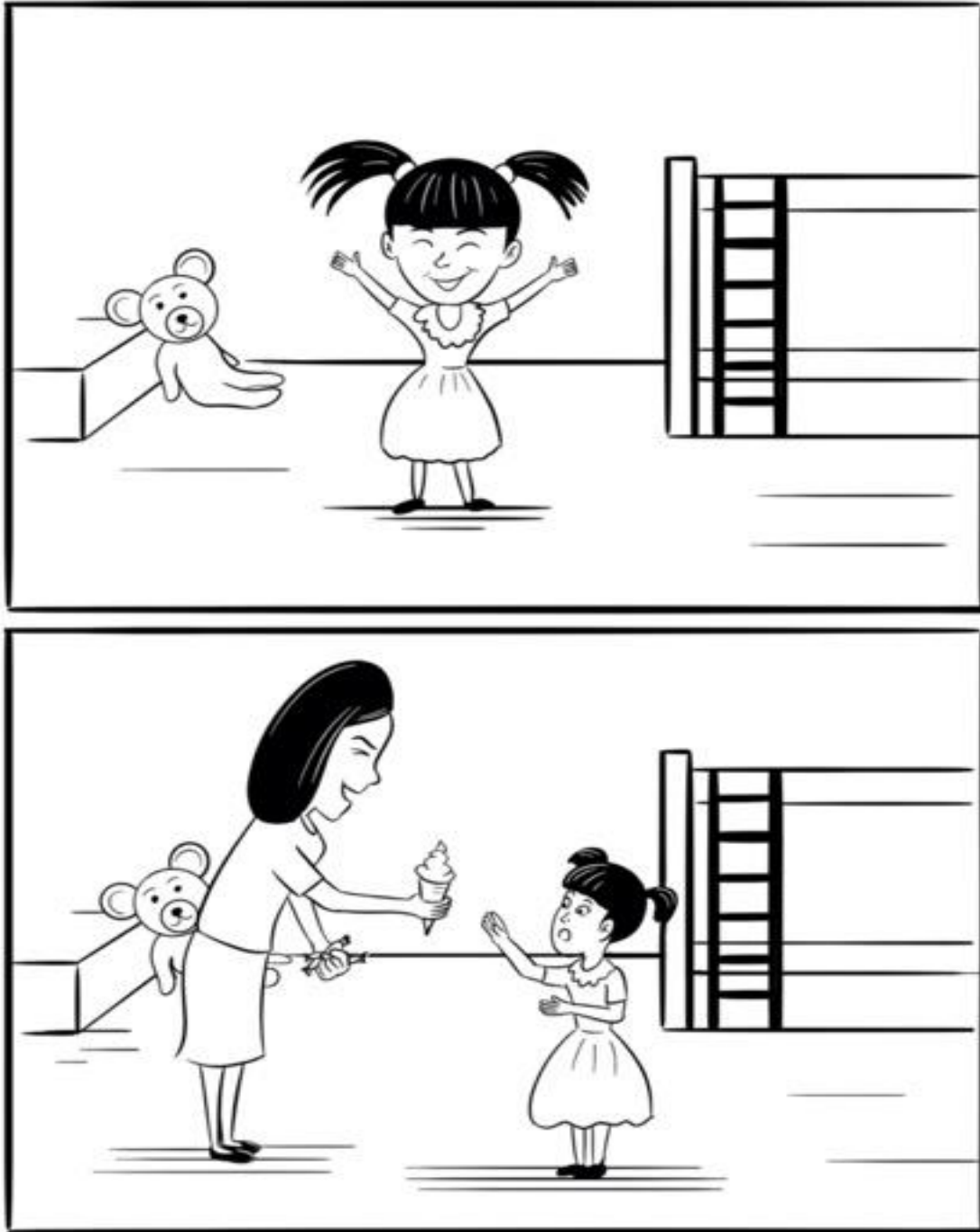


*Figure 6e Manipulation Task Distractor 2: Describe What They Are Doing in Each Picture*





**Figure 7e Cause–Effect Subordination:** Describe the Situation of the Little Girl in the First Picture and Explain Why She is Like That or Why She Feels Like That Based on the Second Picture





*Figure 8e Subjunctive Subordination: What Does the Woman Refuse to Do in Each Picture?*



*Figure 9e Manipulation Task Distractor 3: What Is the Man Doing in Each Picture?*



**Figure 10e** Cause–Effect Subordination: Describe the Situation of the Woman in the First Picture and Explain Why She Is Like That or Why She Feels Like That Based on the Second Picture



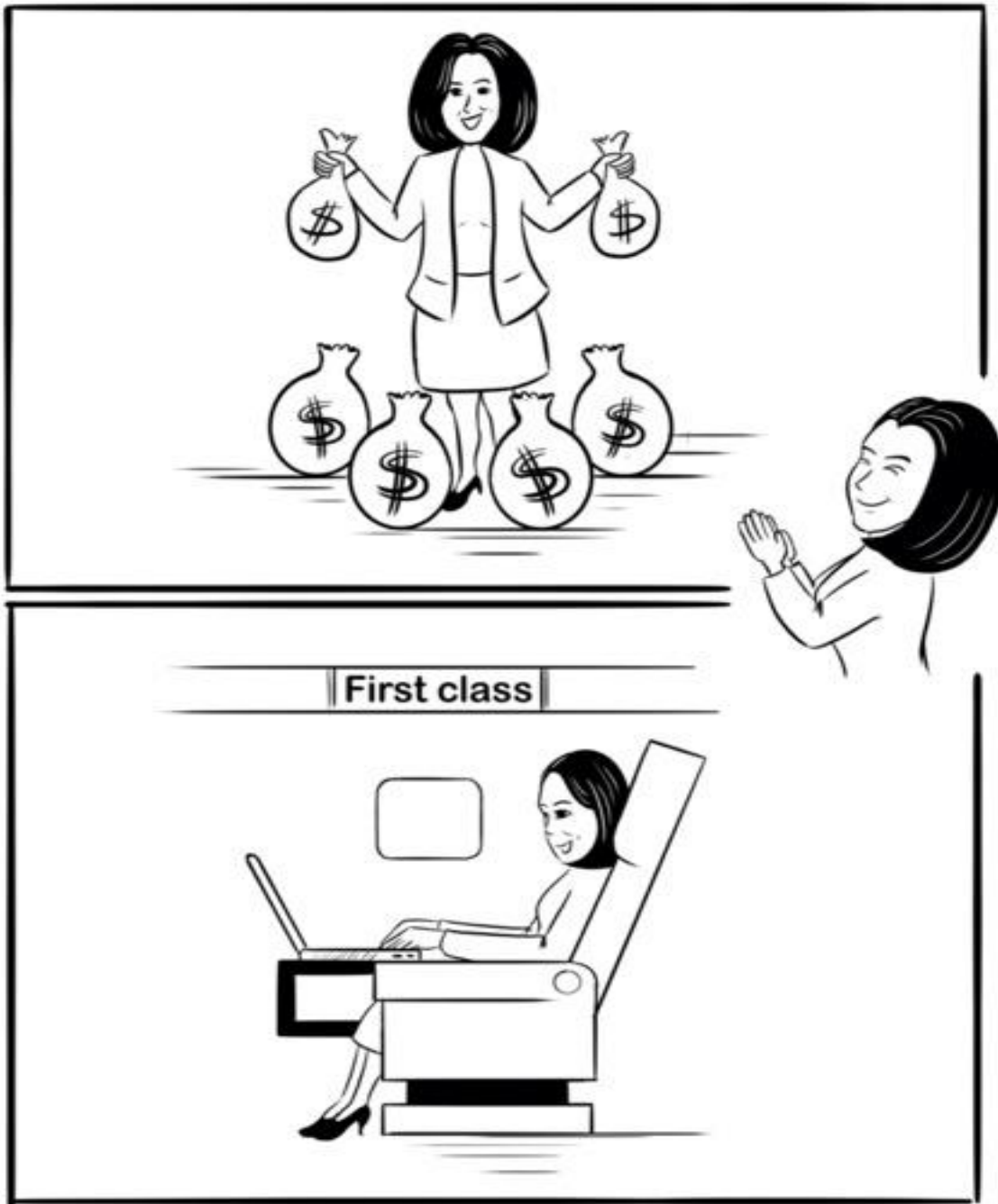
*Figure 11e Subjunctive Subordination: What Does the Man Want to Do in Each Picture?*



**Figure 12e** Cause–Effect Subordination: Describe the Situation of the Woman in the First Picture and Explain Why She Is Like That or Why She Is in That Situation Based on the Second Picture



*Figure 13e Subjunctive Subordination: What Does the Woman Wish to Do in Each Picture?*



*Figure 14e Manipulation Task Distractor 4: Describe What They Are Doing in Each Picture and Where These Two Actions Are Taking Place*





**Figure 15e** Cause–Effect subordination: Describe What the Little Boy Is Doing and Explain Why He Is Doing That Based on the Second Picture





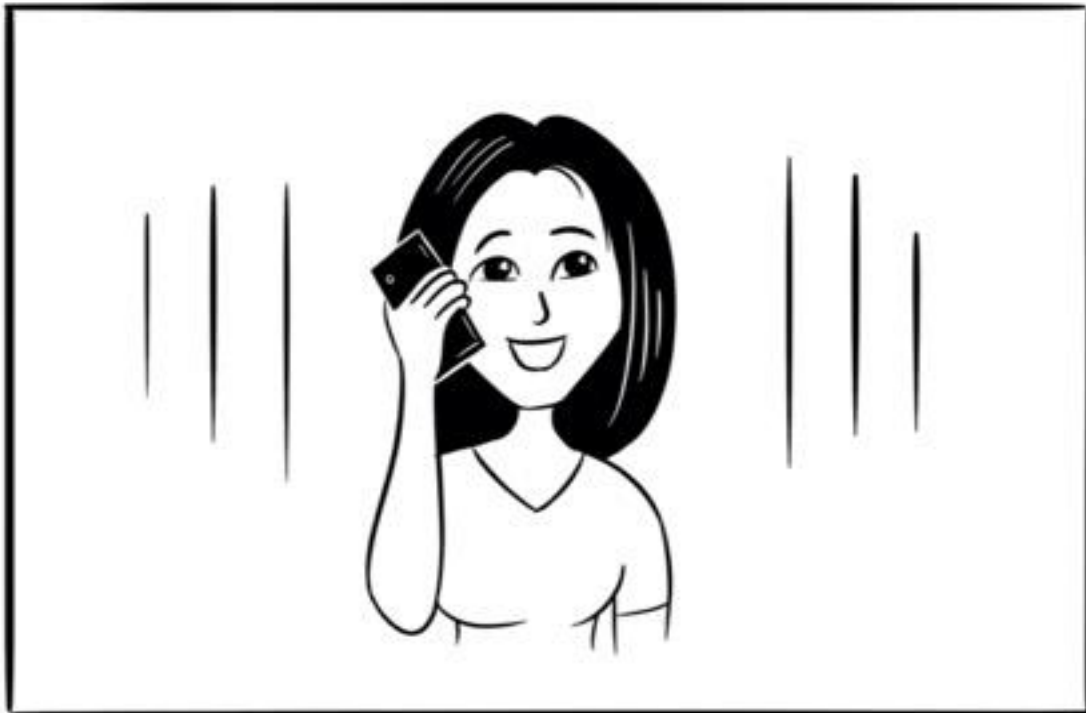
*Figure 16e Subjunctive Subordination: What Is He Trying to Do in Each picture?*



**Figure 17e** Cause–Effect Subordination: Describe Where the Woman Went in the First Picture and Explain Why She Went There Based on the Second Picture.



*Figure 18 e Manipulation Task Distractor 5: Describe What She Is Doing in Each picture*



**Figure 19e** Cause–Effect Subordination: Describe What Happened in the First Picture and Explain Why It Happened Based on the Second Picture



**Appendix F** Tables for Follow-up Comparisons on the Fluency Means of the L2 Participants

**Table 1f** Tukey's Comparisons for the Fluency Mean Scores in the Past Narrative Task Based on the L2 Proficiency Levels (First Year, Second Year, and Third year)

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
2nd year (n = 21)	1st year (n = 19)	2.74	2.48	.514	-3.23	8.72
3rd year (n = 17)	1st year (n = 19)	9.37*	2.61	.002	3.07	15.68
	2nd year (n = 21)	6.63*	2.55	.032	.47	12.79

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 2f** Pairwise Comparisons for the Fluency Mean Scores of the Three Proficiency Levels (First year, Second year, and Third year) Within Each L2 Group (IPG and UIPG) in the Past Narrative Task

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
<b>IPG</b>						
2nd year (n = 11)	1st year (n = 9)	2.28	3.52	.519	-4.76	9.34
3rd year (n = 8)	1st year (n = 9)	5.43	3.80	.159	-2.20	13.06
	2nd year (n = 11)	3.15	3.63	.391	-4.15	10.44
<b>UIPG</b>						
2nd year (n = 10)	1st year (n = 10)	2.90	3.50	.411	-4.12	9.92
3rd year (n = 9)	1st year (n = 10)	12.90*	3.59	<.001	5.69	20.12
	2nd year (n = 9)	10.00*	3.59	.008	2.79	17.22

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 3f** Tukey's Comparisons for the Fluency Mean Scores in the Present Narrative Task Based on the L2 Proficiency Levels (First Year, Second Year, and Third year)

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
2nd year (n = 21)	1st year (n = 19)	4.26	2.53	.222	-1.85	10.38
3rd year (n = 17)	1st year (n = 19)	10.38*	2.67	<.001	3.93	16.83
	2nd year (n = 21)	6.12	2.61	.059	-.19	12.42

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 4f** Pairwise Comparisons for the Fluency Mean Scores of the Three Proficiency Levels (First Year, Second Year, and Third Year) Within Each L2 Group (IPG and UIPG) in the Present Narrative Task

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
<b>IPG</b>						
2nd year ( <i>n</i> = 11)	1st year ( <i>n</i> = 9)	4.24	3.60	.244	-2.98	11.47
3rd year ( <i>n</i> = 8)	1st year ( <i>n</i> = 9)	6.71	3.89	.091	-1.10	14.52
	2nd year ( <i>n</i> = 11)	4.47	3.72	.510	-5.00	9.93
<b>UIPG</b>						
2nd year ( <i>n</i> = 10)	1st year ( <i>n</i> = 10)	3.90	3.58	.281	-3.29	11.09
3rd year ( <i>n</i> = 9)	1st year ( <i>n</i> = 10)	13.67*	3.68	<.001	6.28	21.05
	2nd year ( <i>n</i> = 9)	9.77*	3.68	.011	2.38	17.15

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Appendix G** Tables for Follow-up Comparisons on the Accuracy of the L2 Participants on the Past Tense

**Table 1g** Tukey's Comparisons for the Accuracy on the Past Tense Inflected for the 3.s.m Form Based on the L2 Proficiency Levels (First Year, Second Year, and Third Year)

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
2nd year (n = 21)	1st year (n = 19)	-4.57	8.53	.854	-25.17	16.03
3rd year (n = 17)	1st year (n = 19)	47.25*	9.00	<.001	25.53	68.97
	2nd year (n = 21)	51.82*	8.79	<.001	30.60	73.05

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 2g** Pairwise Comparisons for the Accuracy of the L2 Groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Past Tense Inflected for the 3.s.m Form

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
IPG	UIPG					
1st year (n = 9)	1st year (n = 10)	-37.83*	12.38	.004	-62.69	-12.97
2nd year (n = 11)	2nd year (n = 10)	-28.14*	11.78	.021	-51.78	-4.50
3rd year (n = 8)	3rd year (n = 9)	-12.43	13.09	.347	-38.72	13.87

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 3g** Pairwise Comparisons for the Accuracy of the Three Proficiency Levels (First Year, Second Year, and Third Year) Within Each L2 Group (IPG and UIPG) on the Past Tense Inflected for the 3.s.m Form

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
IPG						
2nd year (n = 11)	1st year (n = 9)	1.94	12.11	.873	-22.38	26.26
3rd year (n = 8)	1st year (n = 9)	60.59*	13.09	<.001	34.30	86.88
	2nd year (n = 11)	58.65*	12.52	<.001	33.51	83.79
UIPG						
2nd year (n = 10)	1st year (n = 10)	-7.75	12.05	.523	-31.95	16.45
3rd year (n = 9)	1st year (n = 10)	35.18*	12.38	.006	10.32	60.04
	2nd year (n = 9)	42.93*	12.38	.001	18.07	67.79

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 4g** Tukey's Comparisons for the Accuracy on the Past Tense Inflected for the 3.s.f Form Based on the L2 Proficiency Levels (First Year, Second Year, and Third Year)

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
2nd year (n = 21)	1st year (n = 19)	2.25	10.89	.977	-24.03	28.53
3rd year (n = 17)	1st year (n = 19)	32.01*	11.48	.020	4.31	59.72
	2nd year (n = 21)	29.77*	11.22	.028	2.69	56.84

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 5g** Pairwise Comparisons for the Accuracy of the L2 Groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Past Tense Inflected for the 3.s.f Form

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
IPG	UIPG					
1st year (n = 9)	1st year (n = 10)	-36.68*	15.80	.024	-68.39	-4.97
2nd year (n = 11)	2nd year (n = 10)	-37.58*	15.02	.016	-67.74	-7.42
3rd year (n = 8)	3rd year (n = 9)	-19.44	16.71	.250	-52.98	14.10

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 6g** Pairwise Comparisons for the Accuracy of the Three Proficiency Levels (First Year, Second Year, and Third Year) Within Each L2 Group (IPG and UIPG) on the Past Tense Inflected for the 3.s.f Form

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
IPG						
2nd year (n = 11)	1st year (n = 9)	3.66	15.45	.814	-27.36	34.68
3rd year (n = 8)	1st year (n = 9)	41.03*	16.71	.018	7.49	74.57
	2nd year (n = 11)	37.37*	15.98	.023	5.30	69.44
UIPG						
2nd year (n = 10)	1st year (n = 10)	4.56	15.38	.768	-26.31	35.43
3rd year (n = 9)	1st year (n = 10)	23.79	15.80	.138	-7.92	55.50
	2nd year (n = 9)	19.23	15.80	.229	-12.48	50.94

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.



**Appendix H** Tables for Follow-up Comparisons on the Accuracy of the L2 Participants on the Present Tense

**Table 1h** Tukey's Comparisons for the Accuracy on the Present Tense Inflected for the 3.s.m Form Based on the L2 Proficiency Levels (First year, Second Year, and Third Year)

Variable (I)	Variable (J)	MD		Sig.	95% Confidence Interval	
		(I vs. J)	SE		Lower Bound	Upper Bound
2nd year (n = 21)	1st year (n = 19)	-6.03	7.08	.673	-23.13	11.06
3rd year (n = 17)	1st year (n = 19)	20.34*	7.47	.024	2.31	38.36
	2nd year (n = 21)	26.37*	7.30	.002	8.76	43.99

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 2h** Pairwise Comparisons for the Accuracy of the L2 Groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Present Tense Inflected for the 3.s.m Form

Variable (I)	Variable (J)	MD		Sig.	95% Confidence Interval	
		(I vs. J)	SE		Lower Bound	Upper Bound
IPG	UIPG					
1st year (n = 9)	1st year (n = 10)	-.30	10.28	.976	-20.94	20.33
2nd year (n = 11)	2nd year (n = 10)	-16.67	9.77	.094	-36.28	2.96
3rd year (n = 8)	3rd year (n = 9)	-23.62*	10.87	.034	-45.44	-1.80

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 3h** Pairwise Comparisons for the Accuracy of the Three Proficiency Levels (First year, Second year, and Third Year) Within Each L2 Group (IPG and UIPG) on the Present Tense Inflected for the 3.s.m Form

Variable (I)	Variable (J)	MD		Sig.	95% Confidence Interval	
		(I vs. J)	SE		Lower Bound	Upper Bound
IPG						
2nd year (n = 11)	1st year (n = 9)	-13.81	10.05	.176	-33.99	6.37
3rd year (n = 8)	1st year (n = 9)	7.99	10.87	.465	-13.83	29.81
	2nd year (n = 11)	21.81*	10.39	.041	.940	42.67
UIPG						
2nd year (n = 10)	1st year (n = 10)	2.55	10.00	.800	-17.53	22.63
3rd year (n = 9)	1st year (n = 10)	31.31*	10.28	.004	10.68	51.94
	2nd year (n = 9)	28.76*	10.28	.007	8.13	49.39

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 4h** Tukey's Comparisons for the Accuracy on the Present Tense Inflected for the 3.s.f Form Based on the L2 Proficiency Levels (First year, Second Year, and Third Year)

Variable (I)	Variable (J)	MD			95% Confidence Interval	
		(I vs. J)	SE	Sig.	Lower Bound	Upper Bound
2nd year (n = 21)	1st year (n = 19)	-1.86	7.84	.970	-20.79	17.07
3rd year (n = 17)	1st year (n = 19)	19.88	8.27	.051	-.08	39.84
	2nd year (n = 21)	21.74*	8.08	.026	2.23	41.25

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 5h** Pairwise Comparisons for the Accuracy of the L2 Groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Present Tense Inflected for the 3.s.f Form

Variable (I)	Variable (J)	MD			95% Confidence Interval	
		(I vs. J)	SE	Sig.	Lower Bound	Upper Bound
IPG	UIPG					
1st year (n = 9)	1st year (n = 10)	9.36	11.38	.415	-13.49	32.21
2nd year (n = 11)	2nd year (n = 10)	-14.44	10.82	.188	-36.16	7.29
3rd year (n = 8)	3rd year (n = 9)	-26.59*	12.04	.032	-50.75	-2.43

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 6h** Pairwise Comparisons for the Accuracy of the Three Proficiency Levels (First year, Second year, and Third Year) Within Each L2 Group (IPG and UIPG) on the Present Tense Inflected for the 3.s.f Form

Variable (I)	Variable (J)	MD			95% Confidence Interval	
		(I vs. J)	SE	Sig.	Lower Bound	Upper Bound
IPG						
2nd year (n = 11)	1st year (n = 9)	-13.66	11.13	.226	-36.01	8.69
3rd year (n = 8)	1st year (n = 9)	.88	12.04	.942	-23.28	25.04
	2nd year (n = 11)	14.54	11.51	.212	-8.57	37.64
UIPG						
2nd year (n = 10)	1st year (n = 10)	10.14	11.08	.364	-12.10	32.38
3rd year (n = 9)	1st year (n = 10)	36.83*	11.38	.002	13.98	59.68
	2nd year (n = 9)	26.69*	11.38	.023	3.84	49.54

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Appendix I** Tables for Follow-up Comparisons on the Accuracy of the L2 Participants on the Past and Present Negation

**Table 1i** Pairwise Comparisons for the Accuracy of the L2 groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Past Negation

Variable (I)	Variable (J)	MD		Sig.	95% Confidence Interval	
		(I vs. J)	SE		Lower Bound	Upper Bound
IPG	UIPG					
1st year (n = 9)	1st year (n = 10)	-38.53*	14.41	.010	-67.47	-9.60
2nd year (n = 11)	2nd year (n = 10)	-15.55	13.70	.262	-43.06	11.97
3rd year (n = 8)	3rd year (n = 9)	-34.04*	15.24	.030	-64.63	-3.44

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 2i** Tukey's Comparisons for the Accuracy on the Verbal Present Negation Based on the L2 Proficiency Levels (First Year, Second Year, and Third Year)

Variable (I)	Variable (J)	MD		Sig.	95% Confidence Interval	
		(I vs. J)	SE		Lower Bound	Upper Bound
2nd year (n = 21)	1st year (n = 19)	-10.78	8.70	.436	-31.79	10.22
3rd year (n = 17)	1st year (n = 19)	22.14*	9.17	.050	-.01	44.28
	2nd year (n = 21)	32.92*	8.97	.002	11.28	54.56

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 3i** Pairwise Comparisons for the Accuracy of the L2 groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Verbal Present Negation

Variable (I)	Variable (J)	MD		Sig.	95% Confidence Interval	
		(I vs. J)	SE		Lower Bound	Upper Bound
IPG	UIPG					
1st year (n = 9)	1st year (n = 10)	-7.40	12.63	.560	-32.75	17.95
2nd year (n = 11)	2nd year (n = 10)	-30.08*	12.01	.016	-54.17	-5.96
3rd year (n = 8)	3rd year (n = 9)	-25.48	13.35	.062	-52.28	1.33

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 4i** Pairwise Comparisons for Accuracy of the Proficiency Levels (First Year, Second Year, and Third Year) Within Each L2 Group (IPG and UIPG) on the Verbal Present Negation

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
<b>IPG</b>						
2nd year ( <i>n</i> = 11)	1st year ( <i>n</i> = 9)	-21.21	12.35	.092	-46.00	3.59
3rd year ( <i>n</i> = 8)	1st year ( <i>n</i> = 9)	12.55	13.35	.352	-14.26	39.36
	2nd year ( <i>n</i> = 11)	33.75*	12.77	.011	8.12	59.39
<b>UIPG</b>						
2nd year ( <i>n</i> = 10)	1st year ( <i>n</i> = 10)	1.46	12.29	.906	-23.21	26.13
3rd year ( <i>n</i> = 9)	1st year ( <i>n</i> = 10)	30.62*	12.63	.019	5.27	55.97
	2nd year ( <i>n</i> = 9)	29.16*	12.63	.025	3.81	54.51

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Appendix J** Tables for Follow-up Comparisons on the Means of the L2 Participants on the Subjunctive and Cause-effect Subordinations

**Table 1j** Tukey's Comparisons for the Mean Scores of the L2 participants on the Subjunctive Subordination Based on the L2 Proficiency Levels (First Year, Second Year, and Third Year)

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
2nd year (n = 21)	1st year (n = 19)	4.85	6.22	.718	-10.18	19.88
3rd year (n = 17)	1st year (n = 19)	33.01*	6.56	<.001	17.17	48.86
	2nd year (n = 21)	28.16*	6.41	<.001	12.68	43.65

Note. MD = Mean of Difference; SE = Standard Error; Sig. = p value; \* = significant.

**Table 2j** Pairwise Comparisons for the Mean Scores of the L2 Groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Subjunctive Subordination

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
IPG	UIPG					
1st year (n = 9)	1st year (n = 10)	-5.62	9.03	.537	-23.75	12.52
2nd year (n = 11)	2nd year (n = 10)	-22.04*	8.59	.013	-39.28	-4.79
3rd year (n = 8)	3rd year (n = 9)	-41.37*	9.55	<.001	-60.55	-22.19

Note. MD = Mean of Difference; SE = Standard Error; Sig. = p value; \* = significant.

**Table 3j** Pairwise Comparisons for the Mean Scores of the Proficiency levels (First Year, Second Year, and Third Year) Within Each L2 Group (IPG and UIPG) on the Subjunctive Subordination

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
IPG						
2nd year (n = 11)	1st year (n = 9)	-2.69	8.84	.762	-20.43	15.05
3rd year (n = 8)	1st year (n = 9)	14.07	9.55	.147	-5.11	33.25
	2nd year (n = 11)	16.76	9.14	.072	-1.59	35.09
UIPG						
2nd year (n = 10)	1st year (n = 10)	13.73	8.79	.125	-3.92	31.38
3rd year (n = 9)	1st year (n = 10)	49.82*	9.03	<.001	31.68	67.95
	2nd year (n = 9)	36.09*	9.03	<.001	17.95	54.22

Note. MD = Mean of Difference; SE = Standard Error; Sig. = p value; \* = significant.

**Table 4j** Tukey's Comparisons for the Mean Scores of the L2 Participants on the Cause–Effect Subordination Based on the L2 Proficiency Levels (First Year, Second Year, and Third Year)

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
2nd year (n = 21)	1st year (n = 19)	-2.28	7.38	.949	-20.11	15.54
3rd year (n = 17)	1st year (n = 19)	18.50	7.78	.054	-.29	37.30
	2nd year (n = 21)	20.16*	7.61	.023	2.42	39.15

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 5j** Pairwise Comparisons for the Mean Scores of the L2 Groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year) on the Cause–Effect Subordination

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
IPG	UIPG					
1st year (n = 9)	1st year (n = 10)	-16.09	10.71	.139	-37.60	5.42
2nd year (n = 11)	2nd year (n = 10)	-23.21*	10.19	.027	-43.66	-2.76
3rd year (n = 8)	3rd year (n = 9)	-54.86*	11.33	<.001	-77.60	-32.11

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Table 6j** Pairwise Comparisons for the Mean Scores of the Proficiency Levels (First Year, Second Year, and Third Year) Within Each L2 Group (IPG and UIPG) on the Cause–Effect Subordination

Variable (I)	Variable (J)	MD (I vs. J)	SE	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
IPG						
2nd year (n = 11)	1st year (n = 9)	-4.87	10.48	.644	-25.91	16.17
3rd year (n = 8)	1st year (n = 9)	-2.07	11.33	.856	-24.81	20.68
	2nd year (n = 11)	2.80	10.84	.797	-18.95	24.55
UIPG						
2nd year (n = 10)	1st year (n = 10)	2.25	10.43	.830	-18.69	23.19
3rd year (n = 9)	1st year (n = 10)	36.70*	10.71	.001	15.19	58.20
	2nd year (n = 9)	34.45*	10.71	.002	12.94	55.95

Note. MD = Mean of Difference; SE = Standard Error; Sig. = *p* value; \* = significant.

**Appendix K** Tables for Follow-up Comparisons on the Overall Proficiency Means of the L2 Participants

**Table 1k** Tukey's Comparisons for the Overall Proficiency Mean Scores of the L2 Participants Based on the L2 Proficiency Levels (First Year, Second Year, and Third Year)

Variable (I)	Variable (J)	MD		Sig.	95% Confidence Interval	
		(I vs. J)	SE		Lower Bound	Upper Bound
2nd year (n = 21)	1st year (n = 19)	-14.40	52.70	.960	-141.64	112.83
3rd year (n = 17)	1st year (n = 19)	219.86*	55.58	<.001	85.71	354.02
	2nd year (n = 21)	234.27*	54.31	<.001	103.16	365.37

Note. MD = Mean of Difference; SE = Standard Error; Sig. = p value; \* = significant.

**Table 2k** Pairwise Comparisons for the Overall Proficiency Mean Scores of the L2 Groups (IPG and UIPG) at Each Proficiency Level (First Year, Second Year, and Third Year)

Variable (I)	Variable (J)	MD		Sig.	95% Confidence Interval	
		(I vs. J)	SE		Lower Bound	Upper Bound
IPG	UIPG					
1st year (n = 9)	1st year (n = 10)	-126.09	76.49	.105	-279.64	27.47
2nd year (n = 11)	2nd year (n = 10)	-180.94*	72.74	.016	-326.97	-34.92
3rd year (n = 8)	3rd year (n = 9)	-245.22*	80.89	.004	-407.61	-82.83

Note. MD = Mean of Difference; SE = Standard Error; Sig. = p value; \* = significant.

**Table 3k** Pairwise Comparisons for the Overall Proficiency Mean Scores of the Three Proficiency Levels (First Year, Second Year, and Third Year) Within Each L2 Group (IPG and UIPG)

Variable (I)	Variable (J)	MD		Sig.	95% Confidence Interval	
		(I vs. J)	SE		Lower Bound	Upper Bound
IPG						
2nd year (n = 11)	1st year (n = 9)	-34.21	74.82	.649	-184.42	116.01
3rd year (n = 8)	1st year (n = 9)	156.40	80.89	.059	-5.99	318.80
	2nd year (n = 11)	190.61*	77.35	.017	35.32	545.90
UIPG						
2nd year (n = 10)	1st year (n = 10)	20.65	74.45	.783	-128.81	170.11
3rd year (n = 9)	1st year (n = 10)	275.54*	76.49	<.001	121.98	429.09
	2nd year (n = 9)	254.89*	76.49	.002	101.33	408.44

Note. MD = Mean of Difference; SE = Standard Error; Sig. = p value; \* = significant.

## Bibliography

- Abboud, P. (1968). *The teaching of Arabic in the United States: The state of the art*. Washington, DC: Center for Applied Linguistics.
- Abboud, P., & McCarus, E. (1983). *Elementary Modern Standard Arabic*. Cambridge: Cambridge University Press.
- ACTFL (American Council on the Teaching of Foreign Languages). (2012). ACTFL proficiency guidelines: Speaking. Retrieved from [www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012/Arabic](http://www.actfl.org/publications/guidelines-and-manuals/actfl-proficiency-guidelines-2012/Arabic)
- Al Alaslāa, S., & Alhawary, M. (2020). Code choice between Standard Arabic and the Saudi dialect by Saudi Twitter users. *Al-‘Arabiyya*, 53, 1–36.
- Al-Batal, M. (1992). Diglossia proficiency: The need for an alternative approach to teaching. In A. Rouchdy (Ed.), *The Arabic language in America* (pp. 284–304). Detroit, U.S: Wayne State University Press.
- Al-Batal, M. (1995). Issues in the teaching of the productive skills in Arabic. In M. Al-Batal (Ed.), *The teaching of Arabic as a foreign language: Issues and directions* (pp. 115–133). Provo, UT: American Association of Teachers of Arabic.
- Al-Batal, M. (2018). Dialect integration in the Arabic foreign language curriculum: Vision, rationale, and models. In M. Al-Batal (Ed.), *Arabic as one language* (pp. 3–22 ). Washington, DC: Georgetown University Press.
- Al-Batal, M., & Glakes, C. (2018). Dialect integration: Students’ perspectives within an integrated program. In M. Al-Batal (Ed.), *Arabic as one language* (pp. 260–278). Washington, DC: Georgetown University Press.
- Albirini, A. (2011). The sociolinguistic functions of codeswitching between standard Arabic and dialectal Arabic. *Language in Society*, 40(5), 537–562.
- Albirini, A. (2018). Proficiency in standard Arabic and its predictors. In M. Alhawary (Ed.), *The Routledge handbook of Arabic second language acquisition* (pp. 331–361). London and New York: Routledge.
- Albirini, A. (2019). Why Standard Arabic is not a second language for native speakers of Arabic. *Al-‘Arabiyya*, 52, 49–71.



- Albirini, A., & Benmamoun, E. (2015). Factors affecting the retention of sentential negation in heritage Egyptian Arabic. *Bilingualism: Language and Cognition*, 8(3), 470–489.
- Albirini, A., Saadah, E., & Alhawary, M. (2019). L1 and L2 transfer to L3 in L3 and L2 learners of Standard Arabic. *Linguistic Approaches to Bilingualism*, 9, 1–58.
- Al-Hamad, M. (2003). *Morphological and syntactic properties in the acquisition of Arabic as a second language: implications for the theory of SLA and for language teaching* [Doctoral dissertation, University of Essex].
- Alhawary, M. (2005). L2 acquisition of Arabic morphosyntactic features: Temporary or permanent impairment? In M. Alhawary & E. Benmamoun (Eds.), *Perspectives on Arabic linguistics XVII-XVIII* (pp. 273–311). Amsterdam, The Netherlands: John Benjamins.
- Alhawary, M. (2009a). *Arabic second language acquisition of morphosyntax*. New Haven and London: Yale University Press.
- Alhawary, M. (2009b). Speech processing prerequisites or L1 transfer: Evidence from English and French L2 learners of Arabic. *Foreign Language Annals*, 42, 367–390.
- Alhawary, M. (2011). *Modern Standard Arabic grammar: A learner's guide*. Chichester, West Sussex: Wiley-Blackwell.
- Alhawary, M. (2013). Arabic second language acquisition research and second language teaching: What the teacher, textbook writer, and tester need to know. *Al-'Arabiyya*, 46, 23–35.
- Alhawary, M. (2018). Empirical directions in the future of Arabic second language acquisition and second language pedagogy. In K. Wahba, Z. Taha & L. England (Eds.), *Handbook for Arabic language teaching professionals in the 21st Century, Volume II* (pp. 408–421). London and New York: Routledge.
- Alhawary, M. (2019). *Arabic second language learning and effects of input, transfer, and typology*. Washington, DC: Georgetown University Press.
- Alhawary, M. (2021). Arabic applied linguistics. In K. Ryding & D. Wilmsen (Eds.), *The Cambridge handbook of Arabic linguistics* (pp. 11–31). Cambridge: Cambridge University Press.
- Al Midhwah, A., & Alhawary, M. (2020). Arabic diacritics and their role in facilitating reading speed, accuracy, and comprehension by English L2 learners of Arabic. *The Modern Language Journal*, 104(2), 418–438.
- Alosh, M. (1997). *Learner, text and context in foreign language acquisition: An Arabic perspective*. Columbus, OH: Foreign Language Publications Office.

- Aoun, J., Benmamoun, A., & Choueiri, L. (2010). *The syntax of Arabic*. Cambridge: Cambridge University Press.
- Azaz, M. (2016). Cross-linguistic effects in L2 acquisition: The case of Arabic determiner phrase. *Al-‘Arabiyya*, 49, 1–24.
- Bachman, L., & Clark, L. (1987). The measurement of foreign/second language proficiency. *The ANNALS of the American Academy of Political and Social Science*, 20–33.
- Badawi, E. (1973). *Mustawayāt al-‘Arabīyah al-mu‘āširah fī miṣr: Baḥth fī ‘alāqat al-lughah bi al-ḥaḍārah*. Egypt: Dār al-Ma‘ārif.
- Barlow, M., & Kemmer, S. (2000). *Usage-based models of language*. Stanford: CSLI.
- Belnap, K. (2006). A profile of students of Arabic in U.S universities. In K. Wahba, Z. Taha, & L. England (Eds.), *Handbook for Arabic language teaching professionals in the 21st century* (pp. 169–178). Mahwah, NJ: Lawrence Erlbaum.
- Blanc, H. (1960). Stylistic variation in Arabic: A sample of interdialectal conversation. In C. Ferguson (Ed.), *Contributions to Arabic linguistics* (pp. 81–156). Cambridge, MA: Harvard University Press.
- Brustad, K., Al-Batal, M., & Al-Tonsi, A. (2004). *Al-kitaab fī ta‘allum Al-‘Arabiyya: Part I* (second edition). Washington, DC: Georgetown University Press.
- Brustad, K., Al-Batal, M., & Al-Tonsi, A. (2007). *Al-kitaab fī ta‘allum Al-‘Arabiyya: Part II* (second edition). Washington, DC: Georgetown University Press.
- Brustad, K., Al-Batal, M., & Al-Tonsi, A. (2010). *Alif baa’*: Introduction to Arabic letters and sounds (third edition). Washington, DC: Georgetown University Press.
- Brustad, K., Al-Batal, M., & Al-Tonsi, A. (2011). *Al-kitaab fī ta‘allum Al-‘Arabiyya: Part I* (third edition). Washington, DC: Georgetown University Press.
- Brustad, K., Al-Batal, M., & Al-Tonsi, A. (2013). *Al-kitaab fī ta‘allum Al-‘Arabiyya: Part II* (third edition). Washington, DC: Georgetown University Press.
- Bulté, B., & Housen, A. (2012). Defining and operationalizing L2 complexity. In A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 21–46). Amsterdam: John Benjamins Publishing.
- Bybee, J. (2006). From usage to grammar: The mind’s response to repetition. *Language*, 82(4), 711–733.

- Bybee, J., & Hopper, P. (2001). *Frequency and the emergence of linguistic structure*. Amsterdam, Netherlands: John Benjamins.
- Chakrani, B. (2015). Arabic interdialectal encounters: Investigating the influence of attitudes on language accommodation. *Language and Communication*, 41, 17–27.
- Corder, S. (1967). The significance of learners' errors. *International Review of Applied Linguistics*, 5, 161–170.
- DeKeyser, R. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *Handbook of Second Language Acquisition* (pp. 313–348). Oxford: Blackwell.
- DeKeyser, R., & Larson-Hall, J. (2005) What does the critical period really mean? In J. Kroll & A. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 88–108). Oxford: Oxford University Press.
- de Bot, K., Verspoor, M., & Lowie, W. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition*, 10(1), 7–21.
- de Bot, K., & Larsen-Freeman, D. (2011). Researching second language development from a dynamic systems theory perspective. In M. Verspoor, K. de Bot & W. Lowie (Eds.), *A dynamic approach to second language development, methods and techniques* (pp. 5–23). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Eisele, J. (2018). One path or multiple paths: Munther Younes on the integrated approach to Arabic instruction. *Al- 'Arabiyya*, 51, 1–23.
- Ellis, N. (1998). Emergentism, connectionism and language learning. *Language Learning*, 48(4), 631–664.
- Ellis, N. (2002a). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24(2), 143–188.
- Ellis, N. (2002b). Reflections on frequency effects in language processing. *Studies in Second Language Acquisition*, 24(2), 297–339.
- Ellis, N. (2005). At the interface: Dynamic interactions of implicit and explicit language knowledge. *Studies in Second Language Acquisition*, 27(2), 305–352.
- Ellis, N. (2012). Frequency-based accounts of second language acquisition. In by S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 193–210). New York: Routledge.

- Ellis, N. (2015). Implicit and explicit language learning. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 3–23). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Ellis, N., & Robinson, P. (2008). An introduction to cognitive linguistics, second language acquisition, and language instruction. In P. Robinson & N. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 3–24). London: Routledge.
- Ellis, R. (1994). *The study of second language acquisition*. Oxford: Oxford University Press.
- Ellis, R. (2009a). Implicit and explicit learning, knowledge and instruction. In R. Ellis, S. Loewen, C. Elder, H. Reinders, R. Erlam & J. Philip, J. (Eds.), *Implicit and explicit knowledge in second language learning, resting, and teaching* (pp. 3–25). Buffalo, NY: Multilingual Matters.
- Ellis, R. (2009b). Measuring implicit and explicit knowledge of a second language. R. Ellis, S. Loewen, C. Elder, H. Reinders, R. Erlam & J. Philip, J. (Eds.) *Implicit and explicit knowledge in second language learning, testing, and teaching* (pp. 31–64).. Buffalo, NY: Multilingual Matters.
- Ellis, R., & Barkhuizen, G. (2005). *Analyzing learner language*. Oxford: Oxford University Press.
- Fakhri, A. (1984). The use of communicative strategies in narrative discourse: A case study of Moroccan Arabic as a second language. *Language Learning*, 34(3), 15–37.
- Featherstone, J. (2018). Preparing Arabic teachers for integration: The Edinburgh model. In M. Al-Batal (Ed.), *Arabic as one language* (pp. 54–72). Washington, DC: Georgetown University Press.
- Fassi Fehri, Abdelkader. (1993). *Issues in the structure of Arabic clauses and words*. Boston: Kluwer Academic Publishers.
- Ferguson, C. (1959a). Diglossia. *Word*, 15, 325–340.
- Ferguson, C. (1959b). The Arabic koine. *Language*, 35(4), 616–630.
- Foreign Service Institute. (2020). *Foreign language training: FSI's experience with language learning*. Retrieved from <https://www.state.gov/foreign-language-training/>
- Gass, S., & Mackey, A. (2002). Frequency effects and second language acquisition: A complex picture. *Studies in Second Language Acquisition*, 24(2), 249–260.
- Gass, S., & Selinker, L. (2008). *Second language acquisition: An introductory course*. New York: Routledge.

- Ginther, A., Dimova, S., & Yang, R. (2010). Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. *Language Testing*, 27(3), 379–399.
- Godfroid, A. (2022). Hypotheses about the interface between explicit and implicit knowledge in second language acquisition. In A. Godfroid & H. Hopp (Eds.), *The Routledge handbook of second language acquisition and psycholinguistics* (pp. 294–307). New York and London: Routledge.
- Hashem-Aramouni, E. (2011). *The impact of diglossia on Arabic language instruction in higher Education: Attitudes and experiences of students and instructors in the U.S.* [doctoral dissertation, California State University of Sacramento].
- Hawkins, R. (2001). *Second language syntax*. Malden, MA: Blackwell.
- Hawkins, R. (2005). Revisiting *wh*-movement: The availability of an uninterpretable [wh] feature in interlanguage grammars. In L. Dekydtspotter, D. Sprouse & A. Liljestr nd (Eds.), *Proceedings of the 7th generative approaches to second language acquisition Conference (GASLA 2004)* (pp. 124–137). Somerville, MA: Cascadia.
- Hawkins, R., & Chan, Y. (1997). The partial availability of universal grammar in second language acquisition: The ‘failed functional features hypothesis’. *Second Language Research*, 13, 187–226.
- Hawkins, R., Casillas, G., Hattori, H., Hawthorne, J., Husted, R., Lozano, C., Okamoto, A., Thomas, E., & Yamada, K. (2008). The semantic effects of verb raising and its consequences in second language grammars. In J. Liceras, H. Zobl & H. Goodluck, (Eds.), *The role of formal features in second language acquisition* (pp. 328–351). New York: Lawrence Erlbaum.
- Hwang, S. & Lardiere, D. (2013). Plural-marking in L2 Korean: A feature-based approach. *Second Language Research*, 29, 57–86.
- Hiver, P., Al-Hoorie, A., & Evans, R. (2022). Complex dynamic systems theory in language learning: A scoping review of 25 years of research. *Studies in Second Language Acquisition*, 44(4), 913–941.
- Horish, U. (2008). Tense. In K. Versteegh et al. (Eds.), *Encyclopedia of Arabic language and linguistics* (pp. 454–458). Leiden, Netherlands: Brill.
- Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics*, 30, 461–473.

- Housen, A., Kuiken, F., & Vedder, I. (2012). Complexity, accuracy, and fluency: Definitions, measurement, and research. In A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 1–20). Amsterdam: John Benjamins Publishing.
- Huntley, E. (2018). Preparing students for the future: Integrating dialect and standard into the Arabic foreign language classroom. In M. Al-Batal (Ed.), *Arabic as one language* (pp. 73–89). Washington, DC: Georgetown University Press.
- Husseinali, Gh. (2006). Who is Studying Arabic and why? A survey of Arabic students' orientations at a major university. *Foreign Language Annual*, 39(3), 395–412.
- Ibn Fāris, Abū al-Ḥusayn. (1997). *as-Ṣāhibī fī fiqh al-lughah al-‘Arabīyyah wa masā’ilihā wa sunan al-‘Arab fī kalāmihā*. In A. Basaj (Ed.). Beirut, Lebanon: Dār al-Kutub al-‘Ilmiyyah.
- Ibn Jinnī, Abū al-Faṭḥ. (2000). *Sirr ṣinā‘at al-‘i‘rāb*. In Ḥ. Hināwi (Ed.). Beirut, Lebanon: Dār al-Kutub al-‘Ilmiyyah.
- Isleem, M. (2018). Integrating colloquial Arabic in the classroom: A study of students' and teachers' attitudes and effects. In M. Al-Batal (Ed.), *Arabic as one language* (pp. 240–259). Washington, DC: Georgetown University Press.
- Kim, K., & Godfroid, A. (2019). Should we listen or read? Modality effects in implicit and explicit knowledge. *Modern Language Journal*, 103, 648–664.
- Kormos, J. (2014). *Speech production and second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Krashen, S. (1985). *The input hypothesis: Issues and implications*. New York: Longman.
- Kuntz, P. & Belnap, K. (2001). Beliefs about language learning held by teachers and their students at two Arabic programs abroad. *Al-‘Arabiyya*, 34, 91–113.
- Lardiere, D. (1998). Case and tense in the ‘fossilized’ steady state. *Second Language Research*, 14, 1–26.
- Lardiere, D. (2000). Mapping features to forms in second language acquisition. In J. Archibald (Ed.), *Second language acquisition and linguistic theory* (pp. 102–29). Oxford: Blackwell.
- Lardiere, D. (2008). Feature-assembly in second language acquisition. In J. Liceras, H. Zobl & H. Goodluck (Eds.), *The role of formal features in second language acquisition* (pp. 106–140). Mahwah, NJ: Lawrence Erlbaum.

- Lardiere, D. (2009). Some thoughts on a contrastive analysis of features in second language acquisition. *Second Language Research*, 25, 173–227.
- Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, 18, 141–165.
- Larsen-Freeman, D. (2017). Complexity Theory: The lessons continue. In L. Ortega & Z. Han (Eds.), *Complexity theory and language development: In celebration of Diane Larsen-Freeman* (pp. 11–50). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Leddy-Cecere, T. (2018). Diverse speaker output in the integrated Arabic classroom: Trends and interpretations. In M. Al-Batal (Ed.), *Arabic as one language* (pp. 199–220). Washington, DC: Georgetown University Press.
- Lee, J. (2002). The incidental acquisition of Spanish: Future tense morphology through reading in a second language. *Studies in Second Language Acquisition*, 24, 55–80.
- LeTourneau, M. (2008). Subordination. In K. Versteegh et al. (Eds.), *Encyclopedia of Arabic language and linguistics* (pp. 360–367). Leiden, Netherlands: Brill.
- Lin, C-W. & Alhawary, M. (2018). Frequency and L1 transfer effects for the perception and production of Arabic lexical stress by L1 English and L1 Chinese learners of Arabic as an L2. In M. Alhawary (Ed.), *The Routledge handbook of Arabic second language acquisition* (pp. 9–37). London and New York: Routledge.
- Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego: Academic Press.
- Long, M., & Robinson, P (1998). Focus on form: Theory, research, and practice. In Doughty, Catherine, Williams & Jessica. *Focus on form in classroom second language acquisition* (pp. 15–41). Cambridge, UK: Cambridge University Press.
- Long, M., & Sato, C. (1983). Classroom foreigner talk discourse: Forms and functions of teachers' questions. In H. Seliger & M. Long (Eds.), *Class-room oriented research in second language acquisition* (pp. 268–285). Rowley, MA: Newbury House.
- Lonney, D., & Lusin, N. (2019). Enrollments in languages other than English in United States institutions of higher education, Summer 2016 and Fall 2016: Final Report. *Modern Language Association of America*, New York: 1–90. Retrieved from <https://www.mla.org/content/download/110154/file/2016-Enrollments-Final-Report.pdf>
- Loschky, L. (1994). Comprehensible input and second language acquisition: What is the relationship? *Studies in Second Language Acquisition*, 16(3), 303–323.
- Mahmoud, Y. (1986). Arabic after diglossia. In J. Fishman (Ed.), *The Fergusonian impact* (pp. 239–251). Berlin: Mouton De Gruyter.

- McCarthy, J. (1979). *Formal problems in Semitic phonology and morphology* [Doctoral dissertation, MIT].
- McCarus, E. (1987). The study of Arabic in the United States: A history of its development. *Al-‘Arabiyya*, 20, 13–27.
- McCarus, E. (1992). History of Arabic study in the United States. In A. Rouchdy (Ed.), *The Arabic language in America* (pp. 207–221). Detroit, USA: Wayne State University Press.
- Meiseles, G. (1980). Educated spoken Arabic and the Arabic language continuum. *Archivum Linguisticum*, 11(2), 118–148.
- Mejdell, G. (2021). What is formal spoken Arabic? In K. Ryding & D. Wilmsen (Eds.), *The Cambridge handbook of Arabic linguistics* (pp. 199–214). Cambridge: Cambridge University Press.
- Mitchell, T. (1986). What is educated spoken Arabic? *International Journal of the Sociology of Language*, 61(1), 7–32.
- MLA (Modern Language Association). (2016). Language enrollment database: 1958–2016. Retrieved from <http://bitly.ws/KIK9>
- Nap-Kolhoff, E. (2017). The development of Dutch object-naming constructions in bilingual Turkish-Dutch children receiving low amount of Dutch language input. In J. Evers-Vermeul & E. Tribushinina (Eds.), *Usage-based approaches to language acquisition and language teaching* (pp. 169–190). Boston and Berlin: De Gruyter Inc.
- Nassif, L. (2018). Integrating colloquial Arabic into the Arabic L2 curriculum: An analysis of learner speech. In M. Al-Batal (Ed.), *Arabic as one language* (pp. 175–198). Washington, DC: Georgetown University Press.
- Nation, I., & Macalister, J. (2010). *Language curriculum design*. London & New York: Routledge.
- Norris, J., & Ortega, L. (2001). Does type of instruction make a difference? Substantive findings from a meta-analytic review. *Language Learning*, 51, 157–213.
- Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics*, 30(4), 555–578.
- Palmer, J. (2007). Arabic diglossia: Teaching only the standard variety is a disservice to students. *Arizona Working Papers in Second Language Acquisition and Teaching*, 14, 111–122.
- Palmer, J. (2008). Arabic diglossia: Student perceptions of spoken Arabic after living in the Arabic-speaking world. *Arizona working papers in SLA & Teaching*, 15, 81–95.



- Piske, T., & Young-Scholten, M. (2009). *Input matters in SLA*. Bristol, UK: Multilingual Matters.
- Polat, B., & Kim, Y. (2014). Dynamics of complexity and accuracy: A longitudinal case study of advanced untutored development. *Applied Linguistics*, 35(2), 184–207.
- Raish, M. (2018). Comparing the complexity, accuracy, and fluency of the written Arabic in the production of advanced learners and native speakers. In M. Alhawary (Ed.), *The Routledge handbook of Arabic second language acquisition* (pp. 265–285). London and New York: Routledge.
- Révész, A. (2009). Task complexity, focus on form, and second language development. *Studies in Second Language Acquisition*, 31, 437–470.
- Richards, J., & Rodgers, T. (2001). *Approaches and methods in language teaching*. New York: Cambridge University Press.
- Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition*, 18, 27–67.
- Robinson, P. (2001). Task complexity, cognitive resources, and syllabus design: A triadic framework for examining task influences on SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 187–318). Cambridge: Cambridge University Press.
- Ryding, K. (1991). Proficiency despite diglossia: A new approach for Arabic. *The Modern Language Journal*, 75(2), 212–218.
- Ryding, K. (2006). Teaching Arabic in the United States. In K. Wahba, Z. Taha & L. England (Eds.), *Handbook for Arabic language teaching professionals in the 21st century* (pp. 13–20). New York: Routledge.
- Ryding, K. (2018). Teaching Arabic in the United States II. In K. Wahba, Z. Taha & L. England (Eds.), *Handbook for Arabic language teaching professionals in the 21st century* (pp. 11–19). London and New York: Routledge.
- Sachs, J., Bard, B., & Johnson, M. (1981). Language learning with restricted input: Case studies of two hearing children of deaf parents. *Applied Psycholinguistics*, 2, 33–54.
- Schmid, H. (2017). *Entrenchment and the psychology of language learning : How we reorganize and adapt linguistic knowledge*. Washington, DC: American Psychological Association.
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11, 129–158.

- Schwartz, B. & Sprouse, R. (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, 12, 40–72.
- Schwartz, M., & Causarano, P. (2007). The role of frequency in SLA: An analysis of gerund and infinitives in ESL written discourse. *Arizona Working Papers in SLA & Teaching*, 14, 43–57.
- Shiri, S. (2013). Learners' attitudes toward regional dialects and destination preferences in study abroad. *Foreign Language Annual*, 46, 565–587.
- Shiri, S., & Joukhadar, C. (2018). Arabic diglossia speaking without mixing: Practices and outcomes from a beginning level. In M. Al-Batal (Ed.), *Arabic as one language* (pp. 154–172). Washington, DC: Georgetown University Press.
- Sībawayh, 'Amr Ibn 'Uthmān. (1988). *al-Kitāb*. In A. Hārūn (Ed.). Cairo, Egypt: Maktabt al-Khānjī.
- Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30, 510–532.
- Skehan, P., & Foster, P. (2012). Complexity, accuracy, fluency and lexis in task-based performance: A synthesis of the early research. In A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 199–220). Amsterdam: John Benjamins Publishing.
- Snow, C., Arlman-Rupp, A., Hassing, Y., Jobse, J., Joosten, J., & Yorster, J. (1976). Mothers' speech in three social classes. *Journal of Psycholinguists Research*, 5, 1–20.
- Shormani, M. (2014). Structural ambiguity in Arabic DPs: A minimalist N<sup>0</sup>-to-Spec approach. *Al-Qalm Journal*, 1, 1–36.
- Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, 60(2), 263–308.
- al-Suyūfī, Jalāl al-Dīn. (1998). *al-Muzhir fī 'ulūm al-lughah wa anwā'ihā*. In M. Jād, M. Abū al-Faḍl & A. Muḥammad (Eds.). Beirut, Lebanon: al-Maktabah al-'Aṣriyyah.
- Tabachnick, B., & Fidell, L. (2007). *Using Multivariate statistics*. Boston: Allyn and Bacon.
- al-Tha'ālibī, 'Abd Al-Malik. (2002). *Fiqh al-lughah wa sirr al-'Arabiyyah*. In A. Mahdī (Ed.). Beirut, Lebanon: Dār 'Iḥyā' at-Turāth al-'Arabī.
- Unsworth, S., (2008). Comparing child L2 development with adult L2 development: how to measure L2 proficiency. In E. Gavrusseva & B. Haznedar (Eds.), *Current trends in child second language acquisition* (pp. 301–336). Amsterdam: John Benjamins.

- Vercellotti, M. (2012). *Complexity, accuracy, and fluency as properties of language performance: The development of the multiple subsystems over time and in relation to each other* [Doctoral dissertation, University of Pittsburg].
- Versteegh, K. (2014). *The Arabic language*. Edinburgh: Edinburgh University Press.
- Wāfī, A. (2004). *Fiqh al-lughah*. Cairo: Miṣr li aṭ-Ṭibā‘ah wa an-Nashr.
- Wahba, K. (2006). Arabic language use and the educated language user. In K. Wahba, Z. Taha & L. England (Eds.), *Handbook for Arabic language teaching professionals in the 21st century* (pp. 139–155). New York, NY.
- Wahba, K. (2023). A historical overview of Arabic grammar instruction. In K. Wahba, Z. Taha & M. Giolfo (Eds.), *Teaching and learning Arabic grammar: Theory, practice, and research* (pp. 3–36). New York and London: Routledge.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H. (1998). *Second language development in writing: measures of fluency, accuracy, & complexity*. Hawaii: University of Hawaii Press.
- Younes, M. (1990). An integrated approach to teaching Arabic as a foreign language. *al-‘Arabiyya*, 23, 105–122.
- Younes, M. (2015). *The integrated approach to Arabic instruction*. London and New York: Routledge.
- Zyzik, Z. (2009). The role of input revisited: Nativist versus usage-based models. *L2 Journal*, 1(1), 42–61.