Vowel Blindness: Computer-mediated help options for Arabic EFL learners

by

Reem Alsadoon

M.A. (Linguistics), Imam Mohammed Bin Saud University, 2007
B.A., King Saud University, 2004

Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy.

in the
Department of Linguistics
Faculty of Arts and Social Sciences

SIMON FRASER UNIVERSITY

Fall 2015

Approval

Name:	**Reem Alsadoon**
Degree:	**Doctor of Philosophy (Linguistics)**
Title:	***Vowel blindness: Computer-mediated help options for Arabic EFL learners***
Examining Committee:	**Chair:** Dr. Yue Wang Associate Professor

Dr.Trude Heift
Senior Supervisor
Professor

Dr. Cliff Burgess
Supervisor
Senior Lecturer

Dr. Christian Guilbault
Internal/External Examiner
Associate Professor
Department of French

Dr. Robert Blake
External Examiner
Professor
Department of Linguistics
University of California Davis

Date Defended/Approved: December 7, 2015

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

c. as a co-investigator, collaborator or research assistant in a research project approved in advance,

or

d. as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

update Spring 2010

Abstract

This dissertation investigates the impact of different types of help options, specifically input enhancement and form-focused glosses, on reducing vowel blindness of Arabic EFL learners. *Vowel blindness* is the term commonly used for Arabic ESL/EFL learners' difficulty in decoding English vowels by transferring L1 habits of relying heavily on consonants and giving little attention to vowels.

Two hundred-fifty Saudi Arabian EFL students at a beginner to low-intermediate level participated in a study based around a specially designed piece of online software, VALE (Vowel-Assistant for Arabic Learners of English) which incorporates English vowel training through input enhancement and form-focused glosses implemented in the context of reading tasks. Input enhancement was achieved typographically by highlighting the vowels in target words in yellow. The form-focused glosses were designed to include segment-focused glosses, syllable-focused glosses, or segment-syllable focused glosses. Each of the four types of support was experienced by a separate experimental group, while a control group received no such help. VALE also delivered most of the data gathering instruments of the study which included a background questionnaire, pre-test, post-test, delayed post-test, and attitude questionnaire. Retrospective interviews were also conducted with 40 participants.

Three sets of research questions are asked to address the effect of VALE help options on reducing vowel blindness. The first and second sets address the effect of type of support on treated/targeted words and on untreated/nontargeted words, in three stages: initial effect (pre-test - post-test change), retention effect (post-test - delayed post- test), and overall effect (pre-test - delayed post-test). The third set of research questions explores the impact of VALE on raising participants' awareness of the vowel blindness problem and on their attitudes towards VALE.

The results for the first set of research questions revealed significant decreases in vowel blindness errors in the short term for treated words, with significant differences between the experimental groups and the control group. Yet, a significant re-increase in vowel blindness errors occurred in the longer term but an overall vowel blindness reduction effect was found over the entire period of the study, particularly for the segment help option. The second set of research questions again found a significant decrease in vowel blindness for untreated words in the short-term, similar for all VALE help options. In the longer term, a major loss of retention occurred for all the help groups; nevertheless, a significant change in vowel blindness errors was still found over the entire period of the study for the untreated words, particularly for the input enhancement help option. Finally, the third set of research questions revealed through the interview data that the learners generally perceived their help option treatment as positively impacting their awareness of the vowel blindness problem. Interview data and attitude questionnaire also showed mostly positive attitudes towards the technical design of VALE and segment help obtained the highest number of positive responses.

Keywords: Vowel blindness; help options; glossing; input enhancement; Arabic EFL learners; CALL

To my parents, husband, siblings, and kids

Acknowledgements

I would like first to express a deep sense of appreciation and gratitude to my senior supervisor, Dr. Trude Heift, for her mentorship, inspiration, encouragement, and for the countless hours of reading and discussing the many drafts of my dissertation. I am especially indebted to her for her tremendous help in developing my research skills which have hugely improved from where I started. I am also very thankful to Dr. Cliff Burgess for his valuable comments and insights on the drafts of my dissertation. Thank you to both of you for accommodating my travel plans and family arrangements. Also, many thanks to the other committee members, Dr. Robert Blake and Dr. Christian Guilbault, for their valuable comments and feedback on many parts of the dissertation.

I am also most grateful to Al-Imam Mohammed Bin Saud University, particularly Dr. Mohammed Alyaidab, for making arrangements and allowing me to conduct the experiment on the English foundation classes in the Summer of 2013. Thanks are also due to all the students who participated in this research project and for their commitment to the end of the project.

Many thanks to Ehab Al-agizy for his help in programming VALE and for his constant support in fixing technical issues and editing the content of VALE. Also, thanks are extended to my friend Arwa Alkalf for her assistance with recording the audio files for VALE.

I would like to express my gratitude to my colleagues and friends at Simon Fraser University, especially Heidi Kent and Anne Rimrott for being like a family to me. Thanks, Heidi, for the precious memories, beautiful laughter, and wonderful friendship. Thanks, Anne, for being my sister at times when I really needed one. Both of you made the challenge of studying abroad so much easier than I was expecting.

I am so grateful to Dr. Philip Scholfield for his thorough edits and insights on the statistical analysis and discussion. Your comments and edits have made the revision process of this dissertation much faster and easier than I expected. Thank you to Al-Imam Mohammed Bin Saud University for sponsoring my graduate studies, financial support in the development of VALE, and for giving me this chance to study at Simon Fraser

University. I am also grateful to Simon Fraser University for supporting my graduate studies through graduate fellowships and a President's Ph.D. award. My gratitude is also extended to Dr. Trude Heift for her financial support through several research assistantships.

My deepest appreciation goes to my husband and kids, Abdulrahman, Deema, and Abdulkareem. This work would not have been possible without your endless love and sacrifices. Your smiles have given me all the reasons in the world to work harder and make each and every one of you proud of me. My dear dad, Abduljabber, I cannot find the words to thank you enough for enduring all those long trips to see and support me and my family in Canada. I cannot thank you enough for all the sacrifices you have made just to see the Ph.D. progress. Sweet Mom, Fatimah, I know how hard it was to not see me and my family for years and I know you were trying hard to be patient just to help me reach my goal. Thanks mom for your patience, love, and prayers. I am grateful also to all my siblings for their care, support, and prayers. Special thanks to my sisters, Mai and Norah, for taking care of my children in the summer of 2015 to help me finish the revisions of my dissertation. I am also grateful to my friend Nama Alsultan for your constant emotional support throughout the whole writing process.

Table of Contents

List of Tables

List of Figures

List of Acronyms

CA	Classical Arabic
CALL	Computer-Assisted Language Learning
CMC	Computer-Mediated Communication
CTML	Cognitive Theory of Multimedia Learning
ESL/EFL	English as a Second/Foreign Language
L1	First Language
L2	Second Language
MSA	Modern Standard Arabic
ODH	Orthographic Depth Hypothesis
SAV	Spoken Arabic Vernacular
SLA	Second Language Acquisition
VALE	Vowel-Assistant for Arabic Learners of English

1 Introduction

Anecdotal observations from English as a Second Language (ESL) teachers and empirical evidence from Second Language Acquisition (SLA) research have documented Arabic learners' difficulty in the textual decoding and encoding of English vowels (Hayes-Harb, 2006; Ryan & Meara, 1991; Saigh & Schmitt, 2012; Stein, 2010; Taylor, 2008; Thompson-Panos & Thomas-Ruzic, 1983). This phenomenon, as initially coined by Ryan & Meara (1991), is commonly referred to as vowel blindness. It hypothesizes that Arabic learners of English transfer their first language (L1) habits of decoding and encoding Arabic words by relying heavily on consonants and giving little attention to vowels.

Bowen (2011) asserts that the resulting vowel blindness in encoding is a result of insufficient decoding of the word form. For this reason, the current research focuses on vowel decoding with the ultimate goal of also providing progress towards the encoding abilities of Arabic learners of English as a Foreign Language (EFL). More specifically, this dissertation investigates whether input enhancement and glossing in a computer-assisted language learning (CALL) environment helps Arabic EFL learners overcome the vowel decoding problem and thus compensates for their lack of vowel noticing and knowledge. In order to make vowels more salient, the author of this dissertation designed the CALL program VALE (Vowel-Assistant for Arabic Learners of English), which highlights and annotates words with known meaning but confusing vowel forms. VALE provides vowel-focused glosses, therefore attempting to minimize first language (L1) interference.

This dissertation comprises six chapters:

Chapter 2 presents a literature review of studies that aimed at identifying the problem of vowel blindness in Arabic learners of English. It then discusses the main factors that contribute to the existence of vowel blindness in Arabic ESL/EFL learners:

graphemic representation, phonological transparency, morphological transparency, and Arabic diglossia. The chapter concludes with a discussion of word recognition skills for Arabic ESL/EFL learners by focusing on two processing factors that might influence their word recognition skills: phonological and morphological awareness.

Chapter 3 provides a literature review of help options that have been used in vocabulary learning and reading comprehension. It further describes the theories that lay the foundation for this research including the interactionist approach, input noticing and processing, and dual-coding theory. The discussion then focuses on the two types of help options that are most relevant to the current research: glossing and input enhancement.

Chapter 4 describes the methodology of the research conducted for this dissertation. It states the research questions which are split into three main sets. The first set of research questions investigates the effect on vowel blindness of help options for targeted words. The second set examines the effect on vowel blindness of help options for untreated words. The third set investigates the study participants' awareness of vowel blindness and attitudes towards VALE, the CALL program designed by the author of this dissertation. The chapter then describes the study participants and the mixed-method research design of the investigation. VALE, which is intended to assist Arabic EFL learners in improving English vowel reading with different help options, is presented followed by a description of the methodology for data collection and analysis.

Chapter 5 presents the results for the research questions of this dissertation with respect to the effect on vowel blindness of VALE's help options for targeted words, the effect on vowel blindness of VALE's help options for untreated words, and the study participants' awareness of vowel blindness and attitudes towards VALE. The chapter provides the quantitative and qualitative data with their statistical analyses that examined the impact of the help options of VALE on vowel blindness.

Chapter 6 first provides a brief summary of the dissertation. Then, it discusses the results of the study research questions based on the research topics of each set of research questions: 1) the effect on vowel blindness of help options for targeted words, 2) the transfer effect on vowel blindness of help options to nontargeted words, 3)

attitudes towards VALE and awareness of vowel blindness. Finally, it concludes with a discussion of the shortcomings of the study and the pedagogical implications it offers for ESL/EFL teachers.

2 Vowel blindness

Chapter 2 reviews the studies that identify the existence and nature of vowel blindness in Arabic ESL/EFL learners. It also discusses the possible causes of vowel blindness by examining linguistic differences between Arabic and English. More specifically, it investigates the differences between Arabic and English orthography in four areas: 1) graphemic representation, 2) phonological transparency, 3) morphological transparency, and 4) Arabic diglossia. The final section on word recognition and processing skills examines the role of phonological and morphological awareness with respect to vowel blindness.

2.1 Identifying the problem

Vowel blindness hypothesizes that Arabic learners of English transfer their first language (L1) habits of decoding and encoding Arabic words by relying heavily on consonants and giving little attention to vowels. Table 2.1 provides a list of problems that have been identified by several researchers for both decoding and encoding English words. The examples were collected from Arabic learners of English reading aloud sentences, choosing the correct word form, and spelling words.

What is unique about the vowel blindness phenomenon is that while Arabic learners frequently fail to perceive and produce the correct vowel structure in the target word, they consistently preserve the consonant structure intact. For instance, in the examples displayed in Table 2.1, the learners perceived *boiled* as *build,* retaining the same consonant structure *bld*. However, the vowels were changed incorrectly. Researchers find this problem peculiar because Arabic ESL/EFL learners provide a different word instead of just misspelling the vowels or omitting them (Bowen, 2011; Ryan & Meara, 1991; Stein, 2010).

Table 2.1. List of examples of the vowel blindness phenomenon recognized by other researchers

Researcher	Year	Skill	Examples from Arabic ESL/EFL learners	Target
Ryan and Meara	1991	Encoding	...climb to the top of the **moments**	mountains
			Decimal, is that the word which means sad	dismal
			You **spread** the sheep from the goats	separate
		Decoding	**Dpartment, autmobile, managment, sufficint**	department, automobile, management, sufficient
Alsulaimani	1990	Encoding	I **build** an egg for tea	boiled
			Broad a ship	aboard
			The sound of **base**	bees
			Bury: a kind of fruit	berry
			He **cough** the ball	caught
		Decoding	**Stopped, presented, sir, higher, grill**	Stupid, president, sure, hair, girl
Bowen	2011	Encoding	"If you write the **world** you **sholud chick** for **spealing** because if your have **mestake** you cant' understand whats **main** and you **chich** the **gramer** because **theses** important so you can write a good **pargrafth**."	Word, should, check, spelling, mistake, mean, check, grammar, these, paragraph
Saigh and Schmitt	2012	Decoding/Encoding	**Hobet, evaluation, bidget, message,…**	Habit, evolution, budget, massage,…

The phenomenon of vowel blindness has been the subject of a number of empirical studies in SLA. For instance, the first empirical evidence of vowel blindness was made by Alsulaimani (1990), who investigated the kinds of problems experienced by Arabic ESL readers. Study participants were asked to read aloud English words displayed on a computer screen. The words all contained vowels that he had identified as problematic for Arabic ESL readers. Table 2.2 displays a selection of commonly

made errors from the study. Note that in all instances, the consonantal sequence remained intact while the vowels were confused.

Table 2.2. Vowel reading errors made by Arabic ESL readers in Alsulaimani's study (1990, p. 359-364)

Target	Realization
Biscuit	Basket
Bowl	Ball
Subtle	Stable
Blew	Below
Spade	Speed
Abroad	Aboard
Circuit	Cricket

Alsulaimani (1990) also tested the encoding skills of his study participants by displaying words on a computer screen and then asking them to use the words in their own written sentences or phrases. Table 2.3 provides examples of the spelling errors found:

Table 2.3. Vowel spelling errors made by Arabic ESL learners in Alsulaimani's study (1990, p. 391-407)

Target	Student Answer
Detain	We have to protein this prisoner.
Borrow	to borough this cup
Steering	the car's strewn wheels
Truth	I think it means the trough.
Caught	He cough the ball.

As indicated in Table 2.3, the study participants in Alsulaimani's study also misread the vowels and thus produced different words when they spelled them.

Ryan and Meara (1991) conducted a follow-up study at the University of Swansea in Wales to examine vowel blindness and the role of L1 interference. Arabic ESL learners (N=30) were divided into three groups: 1) 10 Arabic ESL learners, 2) 10 non-Arabic ESL students, and 3) 10 adult L1 English speakers. According to the researchers, the participants' proficiency levels ranged from low to intermediate. They

were asked to complete a lexical decision task on 100 English frequency words, consisting of ten letters each. These words were first displayed on a computer screen for 1 second and then concealed for 2 seconds. After that, the words reappeared either with the correct spelling or with a missing vowel (e.g., *department*, *revoltion*). Study participants had to judge whether or not the word was spelled correctly. The results show that there is a significant group difference where Arabic ESL speakers committed the highest number of errors when compared to native English speakers who performed the best and to non-Arabic ESL learners who scored in between. Based on these results, the researchers elaborate that the Arabic readers relied excessively on consonantal information and little on vowels. The remaining groups were better at attending equally to both consonants and vowels. Moreover, results from the reaction time data revealed that the Arabic learners were the slowest among the three groups in performing the task. The researchers concluded that their data show "very strong support for the view that Arabic speakers have great difficulty in processing English words" (p. 538).

Following the same line of research, Hayes-Harb (2006) carried out a study similar to the one by Ryan and Meara's (1991), but with two modifications. The first modification was to make sure that the choice of vocabulary was appropriate for the participants' proficiency level. The second modification was to delete both consonants and vowels in the second set of stimuli. This was a letter detection task, whereby the participants were required to circle target letters whenever they saw them during their reading of four comprehension passages. The results of their ANOVA analysis of detection accuracy revealed that Arabic readers made the most vowel detection errors. In informal post-experiment reviews, the study participants stated that they were sometimes aware of their overreliance on the consonantal information in their word-recognition processes. As a result of the study, the researcher urges teachers to develop conscious strategies in addressing the differences between Arabic and English vowels, at least in the early stages of learning English.

The studies cited above identified the vowel blindness phenomenon for Arabic speakers. The following research targeted the nature of the problem by comparing vowel

to consonant processing, manipulating consonantal cues to process vowels, and exploring the types of problematic vowels.

2.2 The nature of the problem

One line of research on vowel blindness was directed towards identifying and exploring the role word-level processes play in increasing the difficulty of vowel processing. For instance, Randall and Meara (1988) attempted to identify the nature of difficulties experienced by Arabic speakers. Their goal was to examine Arabic learners' visual processing of Arabic and Roman letters. The participants were advanced L1 Arabic speakers of English in a graduate program in the UK. The task was to identify as quickly as possible whether a target letter that was first presented on a computer screen was subsequently present or absent in a string of five letters. The results were presented through a plotted graph of their reaction time. The graph exhibited the characteristics of a *W* shape, which reflected the participants' eye-fixation during search strategies: looking at the beginning, middle, and end of the word. The study found that Arabic speakers, even in more advanced reading stages, still employed the visual processing habits of their L1. Randall and Meara suggest that this search strategy seems to be "an enduring one which is transferred to tasks involving search arrays of Roman letters" (p. 144).

The researchers gave the same experimental task to English speakers and found that their reaction time was plotted in a *U* shape. This shape indicated that eye fixation occurred at the beginning and the end for word recognition. The Arabic visual search is shaped by their L1 habit of processing words according to the tri-consonantal root information. For the English speakers, studies (e.g., Stein, 2010) have confirmed that the beginning (onset) and the end of the word (coda) provide enough information for processing the word.

In this context, Stein (2010) examined the sensitivity to the consonantal patterns either in onset or coda that usually accompany certain vowels. Study participants were divided into three groups: 1) Arabic ESL readers, 2) ESL readers from other language backgrounds, and 3) L1 English readers. Stein explains that native speakers of English

are found to use the onset and vowel or vowel and coda information to accelerate their identification of the whole word. By using contextual information of preceding and following consonants, native speakers narrow down the range of possible vowels. Study results showed that, in general, non-native readers display less sensitivity to the consonantal context of vowels than English readers. However, the Arabic ESL readers demonstrated a weak sense of the consonantal context, especially for vowel-to-coda associations. In fact, Stein describes this weakness as “anti-sensitivity” due to their lack of knowledge about vowel associations with consonants. The nonnative learners of other language backgrounds also demonstrated a lesser degree of sensitivity. Yet, they were still able to recognize some patterns in the vowel associations with consonants in onset and coda.

In another attempt to shed light on the problem of vowel blindness, Bowen (2011) performed a study to identify vowel error patterns. For this purpose, a database was created with 250 spelling errors made by Arabic EFL learners in Oman and the UAE. Several teachers developed the database from hand-written assignments that were collected over a period of three years. The analysis suggested that the number of visual spellers is almost twice the number of phonetic spellers (37% and 19%, respectively). Visual spellers are those who tend to make spelling errors by the incorrect placement of a correct vowel, for example, *siklls* for *skills*, or *ti* for *it*. On the other hand, the phonetic spellers are those who make errors by using incorrect vowels, for example, *serias* for *serious*. Accordingly, phonetic spelling errors display a deficiency in the phonetic knowledge of the English vowel system.

In a more in-depth analysis of the vowel and consonant errors, Bowen found that spelling mistakes with vowels were the highest, at 89%, as compared to consonants (11%). Bowen asserts that these results confirm the assumptions of the vowel blindness phenomenon. A deeper analysis of the vowel mistakes revealed four patterns of errors: 1) correct vowels but in the wrong position, such as *hostiry* for *history*; 2) extra vowels added, such as *partaner* for *partner*; 3) vowel(s) omitted, such as *schdle* for *schedule*; and 4) other vowel mistakes, which involved mostly the confusion of short vowels like *chick* for *check*. Figure 2.1 shows the distribution of the general patterns of errors in percentages, indicating that most of the errors occurred with placing a correct vowel in

the wrong place (43%), followed by other vowel mistakes (30%), vowel omitted (16%), and extra vowel added (11%).

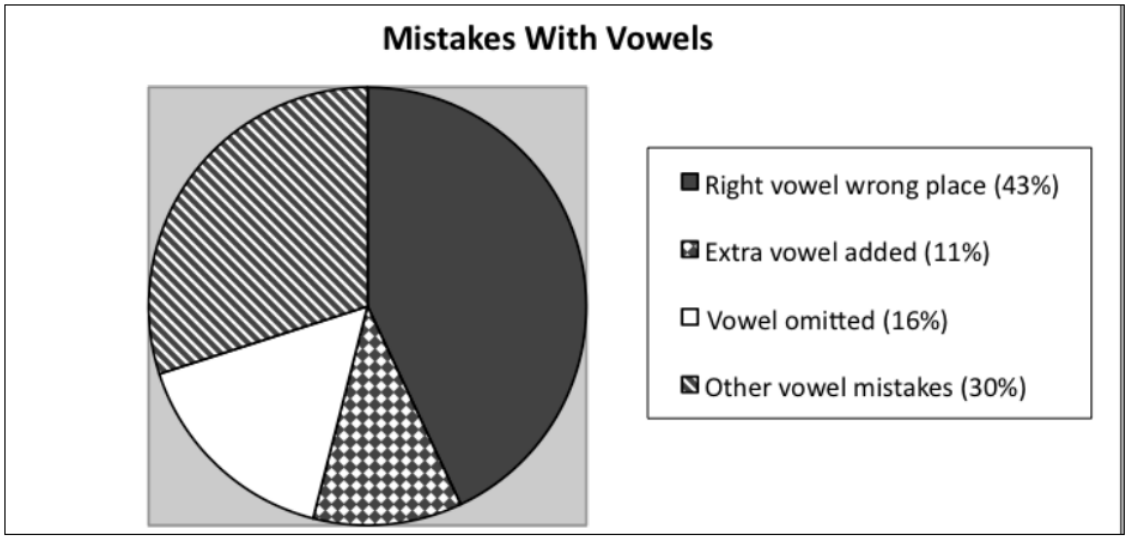


Figure 2.1. General patterns of vowel errors in Bowen's study (2011, p. 92)

With respect to pedagogical implications, Bowen concluded that ESL/EFL teachers need to bear in mind the vowel blindness problem and to find ways to work around it. Bowen encourages techniques such as segmentation, sound-to-letter correspondence, reading aloud, coping, discrimination tasks, and lexical decision tasks to improve vowel processing.

In addition to the research that investigated vowel error patterns, Saigh and Schmitt (2012) were interested in examining vowel blindness with regard to the quantity of vowels: short or long. The study focused on two major sets of vowels: short vowels /ɪ/, /e/, /æ/, and / ʌ/, and long vowels /i:/, /u:/, /a:/, and /ɔ:/. The stimuli consisted of 10 words for each vowel, with a total of 40 words for both set of vowels. Proficiency and familiarity appropriate to the learners' level were considered in the choice of words. The experimental task was designed to test the study participants' ability to notice spelling errors with vowels by asking them to identify spelling mistakes and, if necessary, to correct them. The study affirmed vowels as a problematic area of learning for Arabic ESL learners. However, the short vowels were more difficult in processing and noticing than were the long vowels. Moreover, a deeper analysis of the noticing data revealed

that noticing was harder when the vowel was represented by another vowel but easier when the vowel was omitted, a likely issue for any learner used to an alphabetic system. Yet, and according to Saigh and Schmitt, these findings agree with claims that Arabic ESL learners transfer their vowel processing habits when learning to read and spell in English. As a result, the researchers encourage courseware developers and teachers to employ explicit exercises and means of instruction that enhance noticing and recognition of vowels, especially of short vowels.

Explicit training to improve vowel processing was explored by Taylor (2008) who conducted a training study to investigate the role of direct explicit systematic phonics instruction in improving vowel recognition. The focus of the study was on developing phonological and orthographic awareness to avoid vowel blindness. The subjects were Emirati female students from an English language diploma program at an UAE university. For phonics instruction, the researcher used the "Get Reading" (2006) website. The training period lasted 16 weeks, during which phonics instruction from the "Get Reading" website was introduced differently in two equal phases. In the first phase, the focus was on letters and words using spelling and reading techniques, such as "alphabetizing, teacher-student dictations of short words, and alphabet-to-sound activities" (p. 74). In the second phase, the activities focused more on the sound-to-letter mappings in order to develop phonological and orthographic knowledge of the differences between vowels. Even though the results were not significant with regard to the training effect, the treatment group did exhibit an increase in a post-test of phonological awareness.

The research cited above provides evidence for the nature of the vowel blindness problem, and the difficulties it poses for Arabic learners of English due to their L1 transfer. The following sections discuss possible causes of the vowel blindness phenomenon by considering 1) linguistic differences between Arabic and English, and 2) word recognition and processing skills.

2.3 Linguistic differences between Arabic and English

Researchers such as Bowen (2011), Hayes-Harb (2006), Saigh and Schmitt (2012), and Stein (2010) agree that the vowel blindness phenomenon is attributable to critical differences between the two languages' orthographies. The distinctive nature of graphemic representation, phonological and morphological transparency, and Arabic diglossia creates a disparity between the two systems. These are discussed in the following sections.

2.3.1 Graphemic representation

A major factor in the differences between the Arabic and English systems lies in the graphemic representation of each system. The writing systems of both English and Arabic are based on sound-symbol associations in which phonemes are represented by graphemes. However, English is an alphabetic system whereby graphemes represent phonemes of vowels and consonants. Arabic, on the other hand, is a consonantal system in which graphemes represent only consonants and long vowels (Bassetti & Cook, 2005).

English has a total of 26 letter graphemes. Phonetically, 21 letters represent 24 consonant phonemes, and 5 letters represent 11 vowel phonemes. The English vowel phonemes are provided in Table 2.4.

Table 2.4. English long and short vowels with their spellings adapted from Saigh and Schmitt (2012)

Short vowels	Grapheme representation	Long Vowels	Grapheme representation
/ɪ/	K**i**t, h**y**mn	/iː/	Fl**ee**ce, s**ea**, mach**i**ne
/ɛ/	Dr**e**ss, b**e**d	/uː/	G**oo**se, tw**o**, bl**ue**
/æ/	C**a**t, tr**a**p	/aː/	**A**rticle, p**a**rtner
/ʌ/	L**o**ve, str**u**t	/ɔː/	Rec**o**rd, l**a**w
/ʊ/	F**oo**t, p**u**t	/ɜː/	N**u**rse, st**i**r, c**ou**rage
/ɒ/	L**o**t, w**a**sh		

In contrast, Arabic has 28 letter graphemes representing all 25 consonant phonemes and 3 long vowel phonemes. The three short vowels are represented by diacritics (see Table 2.5). Generally, diacritics appear only at the early stages of teaching literacy to help learners form their early grapheme-phoneme associations and internalize the orthographic patterns of Arabic (Abu-Rabia, 2007).

Table 2.5. Arabic long and short vowels with their spellings adapted from Saigh and Schmitt (2012)

Short vowels	Spelling and status	Long vowels	Spelling and status
/ɪ/	ِ it is not usually marked.	/i:/	ي phonemic status
/æ/	َ it is not usually marked	/a:/	ا or أ phonemic status
/ ʊ/	ُ it is not usually marked	/u:/	و phonemic status

The status of short vowels in Arabic is unique. Native skilled readers can easily understand a text without the short vowel diacritics because in Arabic short vowels convey only grammatical information, such as parts of speech, person, number, case, tense, and voice (Bowen, 2011; Hayes-Harb, 2006). The presence of the diacritics is only helpful for beginning readers, who need to know the grammatical meaning conveyed by the diacritics. For instance, Hayes-Harb (2006) provide an example of this difference, which is illustrated in Figure 2.2 and Figure 2.3 on p. 14.

Figure 2.2 displays the word in the unvowelized form f-ʕ-l. Here, the reader is advanced, or skillful enough to fill in the missing vowel information according to the context. In Figure 2.3, however, the vowel information is presented for beginning readers to assist with their comprehension. Accordingly, the inclusion of short vowels in a text results in a redundancy of syntactic and semantic information that can otherwise be easily inferred from the context. However, because the short vowels lack the visual grapheme-phoneme representation, Arabic ESL readers tend to ignore them.

Brown and Haynes (1985) investigated the impact of the L1 literacy system on visual graphemic processing in reading English. L1 Arabic, Spanish, and Japanese learners who were enrolled in an ESL program at an American University were tested for accuracy and speed in tasks of visual discrimination and matching of words,

pseudowords, and nonsense strings of letters. Study results indicated that the Japanese readers were the most accurate and fastest at decoding words. The Arabic readers were the slowest among all groups. Conversely, the Japanese readers lost their processing advantage when the words became longer, whereas the Arabic readers were more accurate, especially with nonsense strings of words.

Figure 2.2. Unvowelized form of the root f-ʕ-l "the action of doing" (see Hayes-Harb (2006) p. 323)

Figure 2.3. Vowelized form of the root f-ʕ-l: on the left fiʕl: (an act) and on the right faʕala (he did) (see Hayes-Harb (2006) p. 323)

Brown and Haynes concluded that the Japanese readers, with a syllabic and logographic linguistic background, were not able to process words according to the sound-to-letter mapping rules because they were operating in a code system completely different to their L1. On the other hand, the fact that Arabic and Spanish readers were still operating under the same sound-based system (i.e., alphabetic writing system) was an advantage, even though Arabic requires different processing strategies than English. Moreover, the absence of Arabic vowelization as well as the weight of visual processing provides L1 interference due to L1 processing habits.

Fender (2008) discussed L1 influence on ESL spelling and reading in the transfer of grapheme-phoneme correspondence skills. He argued that ESL learners usually rely

on the most familiar letters and sounds before learning new letter-sound associations. Moreover, in a comprehensive review of the studies on developing ESL skills, Figueredo (2006) found that in most studies learners not only transfer their familiarity of letters but also use the same letter-sound mapping patterns.

In addition to the differences in the graphemic representations of Arabic and English, Bowen (2011) illustrated the impact of these differences on other language skills, such as pronunciation, by emphasizing that "vowel blindness is clearly a problem for L1 Arabic learners of English, but 'vowel deafness' should also be recognised. It is the two combined which cause inordinate confusion with vowels, as the errors themselves showed" (p. 95). For instance, many studies investigating perception and production of vowels have documented the difficulty Arabic speakers face in pronouncing vowels (Ali, 2011; Munro, 1992; Nikolova-Simic, 2011; Odisho, 2005). Odisho (2005), for example, reported that pronunciation of vowels is problematic due to Arabic ESL learners employing orthography-based pronunciation in vowel production. Orthography-based pronunciation is observed as a consequence of L1 interference. For example, the English grapheme *a* is realized phonetically as /æ/, /ei/, /ɑ/, /ɛ/, /ɔ/ /ə/ as in <h<u>a</u>t>, <h<u>a</u>te>, <f<u>a</u>ther>, <<u>a</u>ny>, <b<u>a</u>ll>, and <<u>a</u>bout>, respectively (Odisho, 2005). This lack of a one-to-one grapheme-phoneme correspondence in English causes confusion for Arabic speakers. As a result, Arabic learners of English rely on their L1 system and decode the sounds according to the graphemes by, for instance, pronouncing *first* as /fɪrst/ instead of /fərst/.

The major concern here is that if language learners are not able to decode the correct orthographic information of vowels, their internal phonetic representations will be distorted, thereby resulting in misperception and false production of their second or foreign language (L2) speech. In fact, research has shown that speech perception strongly correlates with decoding skills in reading (e.g., Hamada & Koda, 2011) in that a deficiency of utilizing the correct grapheme-phoneme rules in the decoding skills is reflected as a deficiency in speech perception. However, differences in grapheme-phoneme associations and their patterns are not the only source of vowel blindness. The differences between Arabic and English phonology and morphology are additional factors that contribute to this phenomenon. The following sections discuss the

differences in phonological and morphological transparency between the Arabic and English writing systems.

2.3.2 Phonological transparency

Research in reading and spelling has imparted a growing interest in the area of word decoding and encoding in sound-letter (alphabetic) languages (Saiegh-Haddad & Geva, 2008). Transparency of the orthography is commonly tested on the basis of the phonological consistency of the language in question. Katz and Frost (1992) developed the Orthographic Depth Hypothesis (ODH) as the theoretical basis for cross-linguistic studies on writing systems. The ODH posits that languages with a high level of one-to-one sound-to-letter correspondence are described as phonologically shallow and orthographically transparent. On the other hand, languages with a low level of sound-to-letter correspondence, where the same letter or combinations of letters represent different sounds, are considered to have phonological depth and orthographic opacity (Miles, 2000; Stein, 2010; Taylor, 2008).

According to the ODH, English has a deep or opaque orthography and as a result, L2 readers usually require a series of trials to determine the grapheme-phoneme correspondence (Nadia & Charles, 2011). In addition, internalizing these correspondences in order to be able to produce them places another load on the learners' mental processes (Randall, 2009). The opacity of English orthography (as compared to the shallow orthography of such languages as Turkish, Spanish, and Italian) is found to delay even the acquisition of accuracy and automaticity for grapheme-phoneme mapping skills (Aro & Wimmer, 2003). In contrast, Arabic is transparent/ shallow or deep/opaque, depending on the presence and absence of the diacritics of the short vowels. When diacritics are present, Arabic is orthographically shallow, because vowelized Arabic involves less guesswork by providing easier access to phonological forms. When diacritics are absent, however, Arabic is opaque or deep. Only skilled readers are able to decode unvowelized Arabic text with automaticity, fluency, and accuracy. Factors such as the readers' knowledge of the phonology, morphology, and syntax enable them to make correct form-meaning associations through the sentence or the general context (Abu-Rabia, Share, & Mansour, 2003).

A study by Fender (2003), for instance, emphasized that the phonological transparency of English grapheme-phoneme associations constitutes the core problem for Arabic word recognition, especially for vowels. Twenty Japanese and 19 Arabic ESL readers were tested on their ability to recognize words for their phonological mappings with a lexical decision task. Word integration was measured by their ability to attach or detach the target words to phrases and clauses correctly. Fender found that Arabic speakers had more difficulty processing at the phonological word-level than Japanese learners. On the other hand, Japanese learners had more difficulty in the integration of word knowledge into the comprehension of the whole sentence or phrase than did the Arabic group. Fender (2003) argues that the opacity of the English system makes it hard for Arabic learners to find a regular rule for grapheme-phoneme mappings. However, the similarities in phrase and clause structure between English and Arabic, especially when compared to those of Japanese, seem to facilitate integration skills and comprehension.

Apparently, Arabic learners who are accustomed to phonological opacity in orthography are not transferring their habit of filling in missing information from the context accurately. A number of factors revolving around the nature of the English vowel system may influence the accurate reading and spelling of English words. Firstly, although Arabic is opaque when diacritics are absent, regularity, that is, a one-to-one grapheme-phoneme correspondence is still maintained. Readers and writers will still cognitively resort to the same visual representation. In contrast, English is complicated by its irregularity, where the same grapheme can represent more than one phoneme. Secondly, the phonetic variety and complexity of the English vowel system might affect the perception and production of vowels because English, with its lax and tense vowels, introduces several unique vowel qualities, which have no correspondence in the Arabic system. For instance, Naser (1963), Smith (2001), and Saigh and Schmitt (2012) found the following Arabic learner's problems in decoding and encoding English vowels:

- Reliance on phonological information with consistent one-to-one grapheme-to-phoneme processing.
- Overreliance on consonantal information and less noticing of vowels.
- Transfer of the L1 rule of breaking consonant clusters by inserting a vowel in the middle, as the Arabic system does not allow three-consonantal sequencing.

- Problematic encoding and decoding of the nonexistent vowels in the Arabic system such as /e/, /ɔ/, /ɜː/, etc.

In addition to differences in phonology between the two languages, morphology can be another factor in specifying the transparency of a language (Saiegh-Haddad & Geva, 2008), and this is discussed in the following section.

2.3.3 Morphological transparency

The morpheme is the basis of the alphabetic systems of both Arabic and English. According to the Morphological Transparency Hypothesis (Saiegh-Haddad & Geva, 2008), shallow morphological languages such as English are transparent because they usually have stable root and simple derivational processes with suffixes and prefixes. In contrast, deep morphological languages such as Arabic are opaque because they consist of a tri-consonantal root and a series of derivational patterns that use vowels and consonants to create hundreds of variations on the root (Abu-Rabia et al., 2003; Mahfoudhi, Elbeheri, Al-Rashidi, & Everatt, 2010; Saiegh-Haddad & Geva, 2008).

Purnet, Béland, and Idrissi (2000) argue that the root morpheme is an autonomous semantic entity in Arabic and other Semitic languages by which the reader's lexicon is organized in root-based derivations resulting in hundreds of variations from the root. Even though both Arabic and English incorporate derivational morphology to compose multi-syllabic words with root morphemes and affixes, they drastically differ in their morphology from the basic nature of the root to the use of the affixes.

Speaking first of Arabic, its morphology is found to be more complex than English. Arabic is described as consonantal rather than alphabetic mainly because of its tri-consonantal root, which carries the core meaning of most Arabic words. Therefore, the successful decoding of any written word revolves around understanding the derivation process of the word from the root (Mahfoudhi et al., 2010).

The derivation process is the mapping of the root onto different patterns of vowel and consonants affixes to produce a family of words sharing the same core meaning of the root. In other words, Arabic words are made of two abstract morphemes: 1) a root of

three consonants and 2) vocalic and consonantal affixes. The basic meaning comes from the root while the affixes give each word its specific meaning. For example, the root *k-t-b* with its general meaning of "to write" combines with vocalic and consonantal affixes to form ***k**i**t**aa**b*** (book), ***k**aa**t**i**b*** (writer), *ma**k**i**t**ab**a*** (library), and so on.

Ryan and Meara (1991) found the differences between the root and the affixal patterns to be most striking, especially when they occur in texts without vowels. Here, the reader has to fill in the vowel information of different patterns that could be related to the same root. As Ryan and Meara explain:

> In theory, this means that a sentence like "the scribe wrote the book at his desk in the library" could contain five identical k-t-b sequences. In practice, of course, it is more complicated than this. Nonetheless, it is clear that the readers of Arabic are accustomed to a script which places great importance on consonants structures, and plays down the importance of vowels. (p. 533)

On the other hand, the root in English compared to the root in Arabic is always stable, with some prefixes and suffixes used for the derivation of quite a small number of words from the roots. Table 2.6 on p. 20 shows a collection of Arabic and English words derived from the word *write k-t-b*, highlighting the extent of derivation that can possibly occur in Arabic. The Arabic group does not include all the words that are listed in the Arabic dictionary as the list is long for only the main entries of *Lisan Al-Arab Dictionary* by Ibn Manzur (1955) (in this case, around 56 words). However, the English group includes all the words that are listed in the *Oxford Canadian Dictionary*. Accordingly, Table 2.6 indicates the degree of complexity of derivational Arabic, thus making Arabic an orthographically opaque or deep language at the morphological level. In contrast, English is an orthographically shallow language with respect to its morphology.

At first sight, these differences should make English easier to learn for Arabic speakers, as they are accustomed to a more complicated system. However, as Bowen (2011) hypothesizes, the amount of guesswork in the reconstitution of written words in Arabic is not transferable to learning English. Moreover, Fender (2008) argues that although English morphology seems simple compared to Arabic, changes in vowel and stress in the derivational process (e.g., *compete* → *competition* or *competitive)* are confusing and challenging to L1 Arabic learners of English, especially as Arabic plays

down the importance of vowels. Furthermore, some consonants in English change in pronunciation, and this often results in spelling errors from learners who are used to consistent and transparent sound-letter mappings (Ehri, 1997; Fender, 2008). For example, the final *c* in *electric* is pronounced as [k] while in *electricity* it is [s]. Finally, Fender (2008) adds that some derivational morphemes such as *ible* and *able* are confusing in words like *audible* and *dependable* because the only difference is in one vowel. Accordingly, the consistency and degree of transparency of sound-to-letter associations play an important role in learning how to decode English words and especially vowels. The following section discusses how these sound-to-letter associations are established for the colloquial dialects and MSA, the Arabic writing code of formal communication.

Table 2.6. Derived words from the root *k-t-b* and *write*

Tri-consonantal root in Arabic k-t-b	Definition	English root "write"	Definition
Kataba	to write	Write	to write
Kattaba	to make someone write	Writing	the action of writing
Takaataba	to write to each other	Writer	the person who writes
Istaktaba	to dictate	Written	that is done in writing
Kitaab	Book	Writerly	characteristics of a professional author.
Maktab	Office		
Maktaba	library or bookstore		
Kaatib	Clerk		
Miktaab	Typewriter		
Mukaataba	Correspondence		
Mukaatib	correspondent, reporter		
Muktatib	Subscriber		
Kutubii	Bookseller		
Kutayyib	Booklet		

Note. *Sources for the words: Arabic from Lisan Al-Arab Dictionary by Ibn Manzur (1955) and English from "The Oxford Canadian Dictionary of Current English".*

2.3.4 Arabic Diglossia

The phenomenon of diglossia in Arabic is another differentiating characteristic between the Arabic and English writing systems. Diglossia, as Stein (2010) defines it, is "the situation in which many spoken dialects coexist with a standard written literary language. This literary language is not spoken at home, and is first encountered with formal education" (p. 26). Even though English has different spoken dialects, regularity of its orthography is maintained between written and spoken Standard English.

Arabic employs two writing codes: classical Arabic (CA) and MSA (Saiegh-Haddad, 2005). CA is the language of the Holy Book, the Quran, while MSA is the language of the media, formal communication, and education. Spoken Arabic Vernaculars (SAVs) constitute a different group of colloquial dialects according to the geographical area (Saiegh-Haddad, 2005). The differences between the dialects and MSA and CA are phonological, morphological, and syntactic (Abu-Rabia & Taha, 2004).

Acquiring the literacy skills of CA and MSA is similar to learning a second language. For instance, Stein (2010) states that when children start learning CA and MSA, their experience resembles that of L2 learners, based on the fact that they are not spoken at home and first introduced in schools. Some researchers claim that native Arabic speakers do not achieve a true level of literacy in the written Arabic language until they have reached post-high school levels (Saiegh-Haddad, 2005; Taylor, 2008). These claims are supported by Eviatar and Ibrahim (2000), who conducted a meta-linguistic study on bilingual Hebrew-Russian children, monolingual Hebrew children, and monolingual Arabic-speaking children. The study explored language arbitrariness, which refers to the arbitrary assignment of words to referents, such as calling the sun *moon*; phonological awareness; and vocabulary. The study included the following tasks: a game of an arbitrary exchange of words, initial and final phoneme detection, phoneme/syllable deletion, and a vocabulary test. The findings revealed that Arabic-speaking children resemble the bilingual children more than they do the Hebrew monolingual children. Accordingly, Eviatar and Ibrahim (2000) argued that the results indicate that learning classical and MSA Arabic is more like learning a second language, because learners cannot take advantage of the one-to-one correspondence of shallow orthography.

Further evidence of the impact of diglossia on literacy skills in Arabic is provided by Saiegh-Haddad (2005). She investigated the phonological processing of MSA phonemes and vernacular phonemes. The participants were 42 1st grade native Arabic children. They were tested on the basis of five measures of reading processes for pseudowords in MSA and SAV: speed of automatized naming, short-term working memory, phoneme discrimination, phonological isolation, and speed of letter recording. She found that when MSA and pseudowords are encoded into both SAV and MSA, children struggle significantly in analyzing the MSA phonemes. Moreover, first-graders who have just started receiving formal instruction find it hard to isolate MSA phonemes from the SAV phonemes.

In the same context of diglossia, Abu-Rabia and Taha (2004) conducted a study to examine the impact of diglossia in reading and spelling MSA. Arabic speakers were tested on their ability to read and spell texts, isolated words, and pseudowords. The three groups of participants consisted of 20 dyslexic Arabic readers, 20 nondyslexic young readers matching the dyslexic group in their reading level, and 20 readers of the same age as the dyslexic group. The task was to read and spell target words in a text and in isolation (vowelized/non-vowelized) to measure the study participants' accuracy rate as well as the types of errors they make. Results indicate that morphological and semi-phonetic errors rated highest. Semi-phonetic errors are spelling errors based on the mispronunciation of the word. For example, the word ذَهَب/ *went* was read as ذَهَبْ/ *gold* (p. 665). Phonology was also found to be a great challenge in developing spelling skills in Arabic. The researchers emphasize that these difficulties are caused by the diglossic situation in the Arab world because phonology for native speakers is usually developed in the early stages of language acquisition. However, this is not the case with Arabic, where even more advanced students still exhibit these difficulties. The researchers concluded that poor phonological skills may cause an additional cognitive load on working memory, which prevents the development of automaticity in decoding and encoding Arabic words.

Clearly, the studies above reveal the late acquisition of phonological and morphological knowledge of Arabic orthography. Taylor (2008) asserts that the situation of Arabic diglossia extends to the learning of English in that L1 Arabic speakers

experience a limited ability to transfer metacognitive strategies from the L1 to the L2. Thompson-Panos and Thomas-Ruzic (1983) speculate that vowel errors stem from CA in writing and SAV in reading and speaking. Accordingly, diglossia plays an important factor in adding to the disparity between Arabic and English. However, word recognition and processing skills also have an effect on vowel blindness. The following section discusses word recognition and processing skills by examining phonological and morphological awareness.

2.4 Word recognition and processing skills

Word recognition skills are considered to be basic to learning how to decode a text and eventually how to encode it (Hamada & Koda, 2011; Hayes-Harb, 2006). A number of researchers (e.g., Alsulaimani, 1990; Juel, Griffith, & Gough, 1986; Randall, 2009; Ryan & Meara, 1991) argue that word recognition is central to developing reading proficiency. Word recognition involves unitizing a word's meaning with its phonology, morphology, and orthography (word form) (Bowers & Wolf, 1993; Ehri, 1997).

In the literature, the term word recognition is commonly used interchangeably with such terms as word identification, word-level processing, and decoding skills (Frost, 2005; Fukkink, Hulstijn, & Simis, 2005; Hamada & Koda, 2011; Hayes-Harb, 2006; Muljani, Koda, & Moates, 1998). Word recognition involves two related but distinct processes: phonological processing skills and morphological processing skills. Phonological processing is usually the cornerstone for successful word recognition skills (Taylor, 2008). Several researchers (e.g., Abu-Rabia, 2007; Hayes-Harb, 2006; Purnet et al., 2000) argue that morphological processing is equally important, especially for language learners with Semitic native languages and/or Arabic and Hebrew with a root-based lexicon as their L1 (Abu-Rabia, 2007; Koda, 2007; Ryan & Meara, 1991). In fact, Koda (2007) argues that writing systems as a means of communication vary in their conventions and methods of decoding meaning visually. Theoretically, these variations imply more considerations for L2 reading.

The following sections discuss the ways in which the variation in language transparency influence the learner's word recognition by considering phonological and morphological awareness.

2.4.1 Phonological awareness

Phonological awareness is the process of accessing, sorting, identifying, and manipulating phonological information such as phonemes, onset-rhymes, patterns of speech sounds, and syllables (Koda, 2007; Saiz, 2007). There is consensus in the literature about the importance of phonological awareness in learning to read (Hamada & Koda, 2011; Hulme et al., 2002; Koda, 2007). It establishes the foundation for well-specified mappings between the phonological and orthographic representation in reading and hence accuracy in form-meaning associations. Stanovich (2000), for instance, found that poor reading skills uniformly correlate with a deficiency in phonological decoding, or phonological awareness.

In spite of the consensus on the fundamental contributions of phonological awareness to reading, there are, however, different perspectives on its more specific contributions. Some researchers (e.g., Goswami & Bryant, 1992; Muljani et al., 1998) argue that it is a fundamental predictor of reading efficiency and speed. K. Nation and Hulme (1997), for instance, believe that phonological awareness is the key element for the entire reading and spelling process of acquisition. In contrast, other researchers insist that phonological awareness, rather than causing or contributing to reading, is a result of developing reading skills (Morais, Cary, Alegeria, & Bertelson, 1979).

More specifically, Abu-Rabia (1995) conducted a study on 143 Arab children to test their working memory, phonological awareness, and word recognition. The tasks were phonological and orthographic multiple choice tests, a word recognition test, and a missing word test for the working memory. The major finding of the study revealed that poor readers exhibit a deficiency at the phonological and semantic levels. This deficiency was associated with the absence of short vowels in the text, thus leading to a large amount of cognitive guesswork. Moreover, Al Mannai and Everatt (2005) adopted experimental tasks similar to those in Abu-Rabia's study. They investigated the role of

phonological awareness in predicting the literacy skills of four Arabic-speaking children in grades one to three. The researchers emphasized that phonological skills in decoding and encoding Arabic words were effective in indicating the sources and the amount of variability in reading and spelling amongst their study participants. The unvowelized texts were found to pose a problem in developing phonological awareness of the variation in Arabic orthographic transparency.

These results are supported by Nadia and Charles (2011), who recruited 237 Arabic-speaking children from kindergarten to third grade. The study participants were tested on phonological awareness, rapid naming, and phonological memory through an elision and a blending task; a non-word repetition task; and a letter, colour, and object naming task. The study revealed that phonological awareness is a predictor of reading fluency, speed, and word decoding with changes in the grain or unit size of phonological decoding as the children progressed to higher grades. From kindergarten to second grade, the shallow (vowelized) orthography of Arabic required the readers to rely on the grapheme-to-phoneme conversion process rather than on the analysis of larger phonological units. The subjects exhibited a gradual decrease from word to syllable to phoneme level, with the syllable level being more processed and automatized than the phoneme level. Even though children normally achieve automaticity in processing small-to-large phonological units by grade three, the study reports that Arabic readers were found to again go through the process of small-to-large phonological processing due to the absence of diacritics for the short vowels. Automaticity for decoding absent phoneme symbols needs to be developed.

2.4.1.1 Hypotheses on phonological awareness

Cross-linguistically, grain or unit size in the area of orthography decoding has also been of interest to researchers (e.g., Goswami, Ziegler, Dalton, & Scneider, 2003; Ziegler, Perry, Jacobs, & Braun, 2001). The grain-size hypothesis posits that the regularity of the orthography of a language not only contributes to defining the relative lexical and phonological routes but, at the same time, also specifies the preferred grain size of units in reading. Goswami et al. (2003) claim that deep orthography readers of English, for example, switch when decoding between small (e.g., phoneme) and large units (e.g., morpheme), while for readers of shallow orthography languages such as

German and Arabic, small-unit decoding is efficient. The findings from Nadia and Charles (2011) support this claim, as Arabic readers demonstrated the same cognitive behavior in English decoding as in Arabic shallow orthography. They look for the small phonological unit, which is a grapheme-phoneme mapping, and decode it according to the one-to-one mapping of Arabic. The consonants are mostly regular for English but the vowels require decoding of small and large units to resolve their irregularity.

In the case of the decoding and encoding of English texts by Arabic ESL learners, Saigh and Schmitt (2012) demonstrated that beginning and intermediate spellers rely on the L1 strategies of one-to-one grapheme-phoneme phonological processes. As the ODH predicts, learners from shallow languages like Arabic find it challenging to read and spell words from deep orthography languages like English. For instance, Milton and Hopkins (2006) notice that in vocabulary-size tests, native Arabic speakers usually fail to recognize words in orthographic form the way they do in spoken form. Milton and Hopkins assume that:

> These learners can be highly cogent in speech even when vocabulary-size tests suggest their knowledge of English is very limited. One explanation is that the vocabulary knowledge these learners have is phonological in form, and the tests we use, presenting words orthographically for recognition, do not allow these learners to show us what they do know. (p. 31)

Milton and Hopkins provide empirical evidence of the relationship between the development of the orthographic and phonological aspects of the word in the lexicon. They assert that Arabic speaking learners tend to retain more phonological than orthographic knowledge in the lexicon. Namely, the knowledge of the grapheme-phoneme mappings is not yet developed for Arabic ESL readers.

Another important hypothesis in this context is the direction of opacity, which is the direction of mapping from spelling-to-sound or sound-to-spelling. Frost (2005) demonstrated that some alphabetic orthographies, such as Spanish, have a regular mapping in both directions, while other alphabetic orthographies, like French and vowelized Hebrew, are regular from spelling to sound. Arabic falls into both categories in that it is considered regular in both directions if it is vowelized but irregular from spelling to sound if unvowelized.

In summary, the above-mentioned hypotheses demonstrate that grain-size, regularity of the orthography, and direction of opacity play an important role in the word recognition process. They are mainly concerned with the influence of phonological processing in two different linguistic codes. These hypotheses determine the nature of processing phonological units and eventually the whole word. However, morphological awareness is another important factor in developing word recognition skills. The following section discusses morphological awareness, its importance, and the studies that explore morphology as a major component in word recognition.

2.4.2 Morphological awareness

Morphological awareness refers to awareness of the structure of words in order to be able to decode a word's morphological constituents (Koda, 2007). A number of studies have shown that morphological awareness is associated with fluency and efficiency in reading (e.g., Abu-Rabia, 2007; Koda, 2007). Moreover, Abu-Rabia (2007) emphasizes that morphology plays an essential role in the decoding of Semitic texts, such as Arabic and Hebrew. For instance, in a study conducted by Arnbark and Elbro (2002), 33 dyslexic students in fourth and fifth grades in Denmark were trained with 36 lessons in morphological awareness. Throughout the training period, they were tested on phoneme discrimination, spelling, and picture naming. The experimental group displayed a significant improvement in reading comprehension and spelling of morphologically complex words. The researchers confirm that morphological knowledge should be developed from the early stages in order to improve the learner's meaning-oriented decoding strategies in reading and spelling.

Researchers (e.g., Abu-Rabia, 2007; Mahfoudhi et al., 2010; Purnet et al., 2000) argue that morphological awareness, whether in learning another language or literacy skills of the L1, is highly important for learners from language backgrounds with a complex morphology, such as Arabic. For example, according to research (Beland & Mimouni, 2001; Bowen, 2011), Arabic learners are found to transfer their morphological processing skills to English and other languages, such as French. In the context of acquiring literacy skills, research in Hebrew (Ravid, 2001) has found that morphology

specifies the sequence of letter acquisition among children as they are introduced to reading in their early stages.

From an empirical point of view, Abu-Rabia (2007) conducted a study on dyslexic as well as nondyslexic Arabic readers in order to examine the role of morphological awareness in facilitating development of literacy skills in Arabic. The 240 participants were recruited from third, sixth, ninth, and twelfth grades. The tasks were spelling and reading isolated words and reading comprehension. The isolated word task was designed to monitor and measure their morphological decomposing of the word's morphemic constituents. Abu-Rabia found that spelling and morphology is a "powerful predictor of reading accuracy and comprehension" (Abu-Rabia, 2007). Furthermore, morphological awareness is found to be essential in deciding the placement of the short-vowel diacritics on or under the correct letters of the words to indicate accurate pronunciation and correct grammatical function.

In addition, a study by Saiegh-Haddad and Geva (2008) examined the role of morphological awareness in reading accuracy and complex word fluency in Arabic and English with 43 bilingual children in Canada. The tasks included an elision of phoneme test, a morphological relatedness task, and a composition task. In the morphological relatedness task, study participants had to group the words that were derived from the same root. For the morphological composition task, the subjects had to choose the words that could be decomposed into smaller constituents. Study results revealed that morphological awareness was a powerful determinant of "derived-word reading fluency in both languages" (p.499). However, morphological awareness in English and Arabic did not correlate with each other. According to these findings, the researchers argue that morphological awareness is a language-specific linguistic skill, and therefore it is established independently for each language of bilingual children.

Overall, although researchers (Abu-Rabia, 2007; Lukatela, Gligorijevic, Kostic, & Turvey, 1980; Marslen-Wilson, Tylor, Waksler, & Older, 1994; Saiegh-Haddad & Geva, 2008) agree on the importance of morphological awareness for word decoding, its importance depends on the type of language orthography (e.g., Semitic languages, shallowness, etc.) as well as on the level of word processing (e.g., semantic, lexical,

derivational, or inflectional). While a lot of these studies involve dyslexic participants, they provide major insights and hypotheses as to what would make Arabic learners dyslexic when learning to read including the complexity of the phonological and morphological awareness of its orthography for L1 readers. Moreover, Stein (2010) emphasized that these studies explain many sides of the vowel blindness phenomenon, mainly the type of problems that can be transferred from their the L1 to the L2. Indeed, Taylor (2008) confirms that several findings and hypotheses inform vowel blindness research, especially those hypotheses that shed light on the processing mechanism of Arabic learners in decoding Arabic orthography (e.g., Abu-Rabia, 2007) and the effect of L1 orthography in processing the L2 (e.g., Arnbark & Elbro, 2002; Saiegh-Haddad & Geva, 2008). Vowel blindness, after all, is a problem of L1 interference in decoding L2 texts.

The following section presents morphological processing hypotheses which are related to lexical access and morphological awareness.

2.4.2.1 Hypotheses on morphological awareness

Researchers have adopted different hypotheses to explain lexical access and morphological processing. Morphological processing hypotheses are based on cognitive theories, particularly the associative and symbolic theories (Abu-Rabia & Awwad, 2004). The associative theory assumes that words are accessed in the lexicon by their full shape while the symbolic theory claims that words are processed by decomposing words into their root and affixes. For instance, Butterworth (1983) and Henderson, Wallis, and Knight (1984) argue for the whole-word representation hypothesis that claims morphological information is part of semantic processing, which does not require a structured knowledge of morphology at the level of lexical processing. Instead, words are lexically retrieved through access to the whole word as a separate entity in the lexicon and not through decomposition of morphemic constituents (Lukatela et al., 1980). In contrast, the fully decomposed representation hypothesis (Taft, 1981) states that words are morphologically structured in the lexicon, which facilitates word decoding at earlier stages of the word-identification processes. According to this hypothesis, derivational and inflectional processes are similar in the sense that they are both involved in accelerating the decoding of words semantically and morphologically.

The partial decomposing theory, also sometimes referred to as the “stem-based hypothesis,” takes the moderate view in this debate by arguing that “the morphemic units are constructed of stems and inflectional affixes” (Abu-Rabia, 2007, p. 324). It specifies when and how derivational morphology, as opposed to inflectional morphology, is decomposed in the word recognition process. Accordingly, it regards inflectional and derivational morphology as distinct processes with regard to their functionality in the word recognition process. Inflectional morphology occurs at the level of integration between lexical and semantic processes, whereas derivational morphology is more related to the internal organization of the lexicon.

From an empirical point of view, Abu-Rabia and Awwad (2004) investigated the partial decomposing theory on the assumption that “roots are lexical entities that can facilitate lexical access to a large cluster of words that derive from them, whereas word patterns are not lexical entities and have no role in access to words assembled by them” (p. 321). Abu-Rabia and Awwad speculate that low-frequency words require morphological decomposing of their morphemic constituents in order to access their meaning. On the other hand, high-frequency words are processed visually and phonologically as a whole regardless of the word morphology. Yet, in their study only the hypothesis about high-frequency words was tested. The subjects were 48 native Arabic skilled readers at the high school level. A priming-mask technique was used in a lexical decision and naming task, in which a front masking pattern preceded the appearance of a prime word by 500 milliseconds (ms). The target word appeared 50 ms after the primer word. Study results were interpreted based on the associative theory because in Arabic, high-frequency words are accessed as separate entities in the lexicon. However, according to the researchers, the results were inconclusive because only the derivation of nouns was tested. The investigators encourage further research that examines their hypothesis about low-frequency words. In addition, and given that Arabic morphology is similar to Hebrew, Frost, Forster, and Deutsch (1997), in an earlier study, tested the stem-based hypothesis with Hebrew morphology and found that nominal and verbal morphology necessitate decomposing the words for root identification.

2.5 Summary

In summary, the vowel blindness phenomenon refers to the inability of Arabic ESL/EFL learners to decode and encode English words accurately because L1 Arabic speakers rely heavily on consonants and give little attention to vowels. Causes for this inability can be found in the linguistic differences between Arabic and English, and word recognition and processing skills.

The differences between Arabic and English can, for instance, be found in their respective graphemic representations, whereby Arabic uses diacritics for short vowels that are only present for more advanced readers. Moreover, phonological and morphological transparency of Arabic and English differ. With respect to phonology, both English and Arabic are considered opaque languages, although Arabic maintains more regularity in grapheme-phoneme associations. Regarding morphology, Arabic is an opaque language whereas English is a shallow language, and this difference creates difficulties for Arabic ESL/EFL learners in that they transfer their L1 strategies of focusing on consonantal information for the root to guess the derivational form and meaning of the word. Finally, Arabic diglossia is another factor that contributes to vowel blindness. The fact that there are many colloquial dialects and two formal written languages in Arabic complicates the learning of literacy skills for both the L1 and L2.

In addition to these linguistic differences between Arabic and English, word recognition and processing skills are factors that contribute to the vowel blindness problem of Arabic ESL/EFL learners. Morphological and phonological awareness of the L2 orthography is necessary in order to minimize L1 interference (see section 2.4.1 and section 2.4.2). Empirical findings (e.g., Abu-Rabia, 1995; Al Mannai & Everatt, 2005; Nadia & Charles, 2011) stress the importance of developing phonological awareness of English orthography in order to operate within the L2 system rather than that of the L1. As for morphological awareness, there is agreement on its role in word recognition for Semitic languages (Abu-Rabia & Awwad, 2004; Arnbark & Elbro, 2002; Ravid, 2001; Saiegh-Haddad & Geva, 2008). Findings from such research indicate that the ability to process a word morphologically is conducive to lexical decoding, that is, making a form-meaning connection.

Several researchers (Bowen, 2011; Saigh & Schmitt, 2012; Stein, 2010; Taylor, 2008) have suggested that inefficient attention to the word form in decoding is found to also affect the encoding process. For this reason, the current study focuses on vowel decoding as a first step towards assisting Arabic learners of English with their vowel blindness problems. VALE, a CALL reading program specifically designed for the investigation of this dissertation, aims to assist Arabic EFL readers with their noticing of English vowels. It provides help options in the form of input enhancement and form-focused glosses. The following chapter presents a literature review of help options by focusing on glossing and input enhancement studies.

3 Help options

This chapter reviews the help option literature for vocabulary learning and reading comprehension. It also discusses the theories that underpin the use of these help options, including the interactionist approach, input noticing and processing, and dual-coding theory. Then, it examines the literature most relevant to the help options used in the study of this dissertation, particularly glosses and input enhancement.

3.1 Help options in CALL

A large body of CALL research has focused on language support or assistance. Such support has been explored under different terms and classified by Cardenas-Claros and Gruba (2009) as "'help facilities' (Grgurović & Hegelheimer, 2007; Pujolà, 2002), 'guidance support features' (Hegelheimer & Tower, 2004), 'help aids' (Cardenas-Claros, 2005; Grace, 2000), and 'multimedia support resources' (Chun, 2001)." (p. 69). However, Cardenas-Claros and Gruba (2009) prefer the term *help options*, as it implies that learners have a sense of autonomy in whether or not they need support.

Help options are commonly provided in the form of multimedia, closed captions, dictionary references, annotations, translations, hints and illustrative comments, transcripts, and glosses. A gloss is defined by Roby (1999) as additional information supplementing a deficiency in the learner's knowledge of the target input (see section 3.3.1). In addition to help options that provide additional information, a number of researchers (e.g., Chapelle & Hegelheimer, 2000; Sharwood Smith, 1993) emphasize the role of help options in promoting language learning through various techniques, including input enhancement, modification, elaboration, or simplification. Input enhancement is defined by Sharwood Smith (1993) as techniques used to enhance the saliency of linguistic features in the target input (see section 3.3.2). All of these techniques are said to assist learners with noticing specific target forms and establishing

and processing form-meaning connections. Brett (1998), for instance, elaborates that language learning entails two cognitive processes of first noticing the L2 form and then meaning negotiation with the text, computer, teacher, or another learner to map word form to meaning. Blake (2009) emphasizes that video and textual interaction with the computer or with other learners (native or non-native) heightened focus on learning language form.

Help options have been designed for different language skills, such as listening, reading, writing, grammar, and vocabulary learning. According to Cardenas-Claros (2011), transcripts, closed-captioning, and translation are commonly used for developing L2 listening skills. Multimedia glosses, illustrative comments, dictionaries, and L1 translations are commonly used for reading and vocabulary learning. Feedback and hints are usually employed for grammar learning. However, among all of these language learning skills, this dissertation focuses on help options for vocabulary and reading comprehension for Arabic EFL learners.

Help options for vocabulary learning and reading have been extensively researched while other language skills such as listening, writing, and grammar have received less attention from CALL researchers (Cardenas-Claros & Gruba, 2009). Moreover, studies that focus on help options for vocabulary learning (see Cardenas-Claros, 2011; Chun, 2011) have contextualized it within a computer-based *reading* environment. Accordingly, researchers (e.g., Cardenas-Claros, 2011; Cardenas-Claros & Gruba, 2009; Chun, 2011) commonly combine vocabulary and reading studies and this dissertation follows the same classification. The next sections discuss first the theories that are most relevant to the area of help options and then present a literature review of the help options studies for reading and vocabulary learning.

3.2 Theoretical foundations

Several SLA theories lay the foundation for the use of help options in L2 learning. The most relevant theories are the interactionist theory (Gass & Mackey, 2006; Long, 1996), the noticing hypothesis (Schmidt, 1990, 1993, 1994, 1995, 2001), input processing (VanPatten, 1996, 2004), the dual-coding theory (Paivio, 1986), and Mayer's

cognitive theory of multimedia (Mayer, 1997) which provides an extension to the dual-coding theory. Cardenas-Claros (2011) and Chapelle (2009) emphasize the role of these theories informing help option research. These theories are discussed in the following sections.

3.2.1 Interactionist theory

The interactionist theory is one of the main SLA theories informing help options in CALL (see Cardenas-Claros, 2011). It stipulates that interaction is the keystone of second language acquisition (Long, 1996). Interaction brings about comprehension of the language input through *negotiation of meaning.* Long explains that negotiation of meaning "facilitates acquisition because it connects input, internal learner capabilities, particularly selective attention, and output in productive ways" (p. 451). Gass and Mackey (2006) further explain that negotiation of meaning occurs when linguistic problems arise in production. Through interaction, learner feedback that contains modified or comprehensible input is provided, which in turn pushes learners to repair their L2 production. The process by which learners reformulate their production is referred to as *modification*. The intent of modifications is to trigger *noticing* of the mismatches between the learner's interlanguage and the target-like forms of the learner's interlocutors. Noticing is an important step for language learning to take place because it promotes input to become *intake*. Intake in SLA generally refers to the stage at which noticed input is further processed and becomes part of the learner's linguistic repertoire (VanPatten, 1989) (see section 3.2.3).

According to Cardenas-Claros (2011), language learners benefit from interaction with help options through receiving explanatory modifications to the L2 input. Thus, It results in bridging the gap between their language production and the target forms, namely establishing form-meaning connections. Heift and Chapelle (2011) pointed out that the interaction between a learner and the computer can be facilitated through different computer help options and tools, such as checking a dictionary, accessing a concordance, or receiving feedback. In a CALL environment, negotiation of meaning then occurs "when learners ask for repetitions, modifications, or elaborations" (Heift & Chapelle, 2011, p. 561). For this reason, Chapelle (2009) urges CALL developers to

consider techniques such as input enhancement and modified interaction with the goal of facilitating the language learning process. More specifically, she recommends to "make key linguistic characteristics salient by highlighting and providing opportunities for repetitions and modifications for particular forms" (p. 745), and "to support modified interaction between the learner and the computer by providing the learner with control over when to request help, modify responses, and get access to repetition and review" (p. 745).

According to Chapelle, the goal of help options, specifically input enhancement and modification, is noticing and negotiation of meaning of target L2 forms. Attracting the learner's attention is central to the interaction process in order to ultimately make the form-meaning connection and internalize it in the learner's interlanguage system. These concepts are grounded in Schmidt's (1990) *noticing hypothesis*, which stresses that noticing a target form is a starting point for acquisition.

3.2.2 The noticing hypothesis

The noticing hypothesis is a central theoretical underpinning in help options studies (Chun, 2011) and, particularly, in input enhancement studies (Winke, 2013). The noticing hypothesis by Schmidt (1990, 1993, 1995) is based on the notion of *consciousness* in language learning. According to Schmidt (1990), "consciousness is commonly equated with awareness" (p.131). Conscious awareness (i.e., attention) entails two levels: noticing at the lower level and understanding at the higher level. Rosa and Leow (2004) emphasize that CALL help options in the form of feedback, explicit instructions, and input enhancement entail awareness at the level of noticing of converting input to intake and at the level of understanding of hypothesis testing and rule formation. Schmidt (1995) further explains that noticing occurs in short-term memory when conscious attention is applied to the input or parts of the input. However, for the input to become intake, further mental processing is required. On the other hand, understanding occurs in long-term memory when a deeper level of processing results in awareness of the input's general principle, rule, and pattern.

The concept of noticing is important to different areas of SLA. For instance, research on input enhancement has been mainly based on the noticing hypothesis and other attention models. Help options and glossing have investigated noticing from a different perspective, including the cognitive empirical approach of Hulstijn (2001), Hulstijn and Laufer (2001), and Hulstijn, Hollander, and Greidanus (1996) where look-up glossing tasks are used to induce high cognitive effort in noticing and processing the L2 words. Hulstijn (2001) emphasizes that “the more a learner pays attention to a word’s morpho-phonological, orthographic, prosodic, semantic, and pragmatic features and to intraword and interword relations, the more likely it is that the new lexical information will be retained” (p. 259). Heift (2013) and Chapelle (2009) note that help options promote noticing of key linguistic features in the L2 form by providing saliency through highlighting or additional help through repetition and modification.

Most SLA researchers agree that noticing is essential for input to become intake (Truscott & Sharwood Smith, 2011). However, several models suggest a role for attention in noticing. For instance, Tomlin and Villa (1994) propose a functional model of attention and input processing. Based on their model, attention comprises three stages: 1) alertness, which indicates readiness to perceive the input; 2) orientation, which refers to directing attention to a specific type of input; and 3) detection, which is where the input is registered cognitively for further learning processes to occur. According to Tomlin and Villa (1994), detection is "the process by which particular exemplars are registered in memory and therefore could be made accessible to whatever the key processes are for learning, such as hypothesis formation and testing" (p.192). Noticing at this stage results in converting input to intake, similar to Schmidt’s (1990, 1995) term of noticing (Robinson, 1995). However, Tomlin and Villa’s (1994) detection does not require awareness, in contrast to Schmidt’s (1990, 1995) noticing. The work of Tomlin and Villa (1994) has made important contributions in SLA, particularly in the area of input enhancement (e.g., Leow, 1998; Simard, 2001).

Robinson (1995) developed an attention and memory model, which is complementary to Schmidt’s (1995) noticing hypothesis but draws on the attention stages of Tomlin and Villa (1994). Robinson (1995) agrees with the noticing hypothesis in that awareness (i.e., attention) is necessary for the input to become intake. However,

he finds both models unsatisfactory as they do not elaborate on the relationship between attention and the short- and long-term memory systems, specifically on how information is encoded and retrieved from both memory systems (Guidi, 2009). Several researchers build on Robinson's attentional model to account for the impact of help options either through glosses (e.g., Al Ghafli, 2011; Guidi, 2009; Hulstijn, 2001) or input enhancement (e.g., Combs, 2005; Goudarzi & Moini, 2012; Rott, 2007) on the retention of the attended target forms in short and/or long-term memory.

Robinson's (1995) model proposes that detection activates the short-term memory system. However, the input needs to be active in the short-term memory through *rehearsal* before it can become input. According to Robinson (1995), the mechanism for learning is thus "detection plus rehearsal in short-term memory, prior to encoding in long term memory" (p. 297). Accordingly, the attention process includes detection, rehearsal, access to awareness, and then possibly encoding in long-term memory. The model also discusses the effect of external factors on attention, such as input enhancement (Sharwood Smith, 1991, 1993). Robinson (1995) emphasizes that input presentation and task design are important in stimulating conceptually-driven processing rather than data-driven processing. Conceptually-driven processing connects the input to prior knowledge in the long-term memory and is thus more cognitively demanding in the retrieval and encoding of new information. On the other hand, data-driven processing is more mechanical, as it directs the learner's attention to a small list of items, resulting in less cognitive processing.

Within the noticing hypothesis, Schmidt (1994, 1995) has emphasized the role of intentionality and explicitness of the linguistic target form in order to be noticed and learned. The intentionality of L2 learning has been the focus of a number of help option studies in glossing (e.g., Bowles, 2004; Chun & Plass, 1996; Hulstijn, 2003) and in input enhancement (e.g., Gascoigne, 2006; Guidi, 2009; Hulstijn, 2001; Rott, 2007). Furthermore, Jung (2015) reported that explicitness in glossing studies was examined at varying degrees, for instance, using a definition or a synonym (e.g., Guidi, 2009), definition and exemplar sentence (e.g., Hulstijn & Laufer, 2001), and L1 translation accompanied with multimedia cues (e.g., Al-Seghayer, 2001).

The term *intentionality* has been used in the literature to refer to intentional/incidental learning and/or knowledge (Hulstijn, 2001; Schmidt, 1994, 1995), whereas *explicitness* refers to explicit/implicit learning and/or knowledge (Hulstijn, 2005; Schmidt, 1995). These terms have various conflicting interpretations and are sometimes confused with each other (Guidi, 2009; Hulstijn, 2005).

For Schmidt (1994, 1995), the distinction between intentional/incidental learning does not fall under the conceptual categories of conscious/unconscious learning. However, a key question is "whether or not the knowledge gained through incidental learning is represented mentally in a different fashion from knowledge gained through intentional approaches to learning" (Schmidt, 1994, p. 6). Evidence from psychology research reveals no differences between the two learning approaches with respect to their mental representation. However, Paradis (1994) argues that knowledge acquired through intentional/incidental learning approaches is mentally represented differently: a) incidental learning leads to implicit knowledge and automatic use, and b) intentional learning leads to explicit knowledge and no automatic use. In any case, Schmidt (1995) claims that intentionality/incidentality in language learning is not as necessary as attention when designing a language learning task. Intentional learning can only be effective if the task is not attentionally attractive to L2 learners. Therefore, Schmidt concludes that "there can indeed be learning with no intention, but this does not imply the existence of unconscious learning in any other sense" (p. 8).

For Hulstijn (2001, 2003), incidental learning, specifically in a research setting, refers to performing a language task that entails processing L2 information without warning learners that they will be tested on this information. On the other hand, intentional learning occurs when the learners are told in advance about being tested on their knowledge of the task information. With respect to the education context, Laufer and Hulstijn (2001) suggest adopting Schmidt's (1994) general definition for incidental learning: learning something without intent, such as learning vocabulary while the main intention is to communicate.

As for explicitness, Schmidt (1994) distinguishes between explicit/implicit learning and knowledge by referring to the learning process for the former and to the

end-products of learning for the latter. In both cases, the notion of explicitness is based on conscious awareness. Furthermore, Schmidt stresses that more distinctions should be made between explicit learning and instruction. He explains that learning may include forming a conscious hypothesis about the roles of the target language without being told about them. Empirically, he recommends distinguishing between explicit instructed learning and input enhancement techniques and their effect on consciousness and awareness.

Hulstijn (2003), however, emphasizes that SLA researchers should not confuse intentional learning with explicit learning or knowledge. Although it is possible that implicit learning can occur incidentally, explicit learning can still occur intentionally and incidentally. Hulstijn (2005) defines explicit learning as processing language input consciously and intentionally to locate regularities in the input. In contrast, implicit learning is when input processing occurs unconsciously without the intention to capture these regularities. In fact, Hulstijn explains that explicit and implicit knowledge is merely a by-product of each type of learning.

Other researchers in the area of input enhancement and modification (e.g., Alanen, 1995; Robinson, 1996; Shook, 1994) focus more on the stimulus rather than the knowledge or the type of learning by employing implicit/explicit learning to refer to the amount of information presented in the input. According to this approach, explicit/implicit learning is used synonymously with incidental/intentional learning.

In summary, conscious awareness is the main goal of all these cognitive approaches through either incidental, intentional, explicit, or implicit language learning. Conscious awareness entails noticing input to then become intake and then understanding intake to become output. These processes are discussed in VanPatten's (1994, 2011) input processing model below.

3.2.3 Input processing

Once input has been noticed by the learner, it must be subjected to higher-level processing in order to be converted to intake. In Schmidt's theory, awareness plays a role in this processing. For VanPatten (2011), however, noticing in itself is not important,

but what matters is connecting what learners notice to its meaning and function. For this reason, SLA research should pay attention to the distinct nature of the two processes (i.e., noticing and processing). He calls for a processing-based approach to language learning, which would train language learners to "attend to form-meaning links by noticing the relevant cues in the utterances" (p. 19). Chapelle (2009) emphasizes the impact of input processing theory in underlying the theoretical principles for the design of help options in CALL to primarily promote noticing and processing of form-meaning mapping.

In his input processing theory, VanPatten (1996) adds two more stages to Schmidt's (1995) noticing hypothesis: accommodation and restructuring. In the accommodation stage, the input is stored in working memory. The restructuring stage is where the deep processing occurs for the input to be stored in long-term memory. From an input processing perspective, VanPatten and Cadierno (1993a) explain three processes involved in L2 learning. The first process involves noticing specific forms in the input which then become intake. The second process entails the further submission of those forms into the learner's developing system. Finally, the third process involves the learner's ability to produce those forms in communication.

This model raises a question about the mechanism that regulates what gets noticed and processed in the input and why. VanPatten (1996, 2011) proposes three major principles that govern the mechanism of input processing.

- Learners process content words in the input before anything else.
- Learners tend to process lexical information before the grammatical form to establish matching lexical forms first.
- Learners process the first noun or pronoun in the input.

Chapelle (2009) notes that these three principles provide the basis for the development of help options whereby noticing should be the main focus to promote form-meaning mapping. She suggests that the materials should be designed to assist learners to primarily attend to meaning, and to be accompanied by exercises that help in attending to linguistic features of the form. Indeed, the input processing principles convey the dominance of the communicative content in any processing operation before

learners make more general associations between morpho-syntactic information and the input.

Han, Park, and Combs (2008) discuss meaning and form processing through two kinds of information processing: simultaneous and sequential processing. Even though these approaches seem distinct, they operate under general principles of information processing theory. According to the authors:

- Input processing is totally selective.
- Simultaneous processing of two types of information can successfully and efficiently occur only when one type of information has already been automatized and, hence, less demanding of conscious attention.
- If two types of information were not automatized, then simultaneous processing would result in inefficient processing for either one or both types of information. Therefore, sequential processing would be more effective in this case.

For this reason, Han et al. (2008) emphasize that it is more effective to design L2 learning tasks that direct the learner's attention to meaning and grammatical forms separately but sequentially. In this context, they found that simultaneous and sequential processing resembles the distinction between incidental and intentional processing. Simultaneous processing is mainly for meaning where processing for the enhanced or modified form is assumed to occur incidentally. On the other hand, sequential processing is first for meaning so that the attentional resources are then freed up for the form processing.

Input processing sheds some light on the cognitive processing mechanism based on noticing and processing of the form and meaning features of the language input. It provides a framework for the design of the help option tasks in language learning. The role of using *multiple* help options in delivering language information on processing and cognition is discussed through dual-coding theory.

3.2.4 Dual-coding theory

The *dual-coding theory* of Paivio (1969, 1986) proposes that there are two reception systems for memory and cognition: verbal and nonverbal. The verbal system

processes language-related information, such as spoken and written words. The non-verbal system, on the other hand, processes mental imagery information, such as visual objects and gestures. These systems require the activation of two sensory modalities in order to function simultaneously yet independently from each other. For better memorization, the theory suggests that information should be provided in both the verbal and non-verbal mode to activate the two systems. For instance, input enhancement of the target form serves to activate visual processing of the written word.

With respect to help options in CALL, Chun and Plass (1996) argue for dual-hypermedia presentations for three reasons: 1) dual route of information retrieval lead to detailed decoding of the input, 2) activating two systems enhances the chances of information storage in at least one system, and 3) individual characteristics and learning styles can be satisfied with dual routes of processing. Similarly, Chapelle (2005) asserts that multiple modes for help options provide more comprehensible input because it is enhanced by the activation of two cognitive processing systems. However, she emphasizes that the combinations of help options should be designed to achieve two major effects in the input: 1) saliency and 2) additional information for unknown linguistic elements.

Mayer (1997) expanded the dual-coding theory by proposing the cognitive theory of multimedia learning (CTML). This theory follows the same reasoning that auditory (verbal) and imagery (non-verbal) channels are used when processing information. According to the CTML, the two reception systems (i.e., verbal and non-verbal) integrate information at some point in the lexicon, which suggests that they are not completely independent from each other. For Mayer, the selection and organization of information occur separately in two different processing systems. However, the information is integrated at the final stage of information perception.

In the case of language learning, the CTML assumes that two verbal systems (L1 and L2) operate separately but have a common visual system. Therefore, the perception of information is enhanced when it is presented both visually and verbally in order to establish a link between the two systems Mayer (1997). Jones and Plass (2002) assert

that dual-hypertext (written) and hypermedia (pictorial) glosses result in a stronger effect on the retention of the L2 input than one type of gloss or no glosses.

The following sections review the help options studies for L2 vocabulary and reading comprehension that use the help options most relevant to this dissertation, namely, input enhancement and glosses.

3.3 Help options for L2 vocabulary and reading comprehension

Many researchers (Abraham, 2007; Al-Seghayer, 2001; Ariew & Ercetin, 2004; Chapelle & Hegelheimer, 2000; Chun, 2001a, 2001b; Chun & Plass, 1996, 1997; Davis & Lyman-Hager, 1997; De Ridder, 2000, 2002; Ercetin, 2003; Gettys, Imhof, & Kautz, 2001; Grace, 1998; Hegelheimer, 1998; Laufer & Hill, 2000; Li, 2010; Lomicka, 1998; Nagata, 1999; Nikolova, 2004; Peters, 2007; Sakar & Ercetin, 2005; Yanguas, 2009; Yoshii, 2006; Yoshii & Flaitz, 2002; Yun, 2011) have explored the impact of different types of hypermedia help options on vocabulary gain, retention, and reading comprehension. The findings from these studies indicate a positive effect of help options on learning outcome. The following sections examine the help options used in the study of this dissertation: glosses and input enhancement.

3.3.1 Glossing in CALL

According to Chun (2011) hypertext/hypermedia glosses are the most common types of help option for L2 vocabulary and reading comprehension. P. Nation (2001) notes that glosses are commonly employed because they help the learner decode the text by providing additional knowledge at the micro and macro level of form and meaning processing.

Roby (1999) defines glosses as “attempts to supply what is perceived to be deficient in a reader’s procedural or declarative knowledge” (p. 96). Other researchers (Al-Seghayer, 2001; Cardenas-Claros & Gruba, 2009; Chapelle, 2003; Hegelheimer, 1998) refer to glossing as input modification because glosses modify the input to be

more accessible to language learners. They describe the process as “the provision of an accessible rendition of the L2 input” (Chapelle, 2003, p. 45). Cardenas-Claros and Gruba (2009) elaborate on the concept of modification by designing tasks with additional visual information, L1 translation, and L2 definitions in the form of annotations (i.e., glosses).

Chapelle (2003) specifies the role of the interactionist framework, namely that of interaction and noticing, in the design principles for multimedia CALL materials. She emphasizes that instructional materials should provide the learner with input enhancement and modification. Chapelle (2001) further elaborates that modified interaction supports the learner in the case of a breakdown in the meaning-making process during reading comprehension. She contends that glosses with grammatical or semantic clarifications reinforce or verify existing or new information.

Empirically, a number of studies show that glossing is an effective tool to foster interaction and increase noticing. For instance, Chun and Plass (1996) used the interactionist framework to examine learners’ incidental learning while interacting with different hypermedia glosses. More specifically, they measured vocabulary acquisition in the interactive incidental reading environment *Cyberbuch*. The participants were 160 L2 German learners at the university level. Their performance was measured by translation and word recognition tasks. The findings revealed that interacting with hypermedia glosses facilitated incidental learning of vocabulary. The rate of incidental learning was higher than expected, with 25% accuracy and 77% in word recognition. The authors suggest further research to investigate the potential of interactive multimedia glossing and L2 reading.

Yanguas (2009) examined the noticing hypothesis by using multimedia glosses in reading. This study investigated the effect of different hypermedia glosses (picture and text, picture only, or text only) on prompting noticing and consequently on reading comprehension and vocabulary learning. Learners (N=94) read a text under one of the hypermedia gloss conditions. Pre- and post-tests of reading comprehension and word recognition were employed, along with think-aloud protocols, which were included to provide a qualitative measure. The quantitative and qualitative analyses revealed that all

the experimental groups benefitted from the hypermedia glosses in noticing and recognizing significantly more target words than the control group. Comprehension test data showed that the combined hypermedia glosses condition led to better performance as compared to other experimental conditions and control.

For vocabulary gain and retention, dictionaries, multimedia annotations of text and pictures, and L1 translation were found to be effective in short-term (immediate effect) and long-term (delayed effect, generally after 4 weeks) learning (Abraham, 2007, 2008; Hegelheimer, 1998; Karp, 2002; Laufer & Hill, 2000; Li, 2010; Nikolova, 2004). For example, Johnson and Heffernan (2006) examined vocabulary retention in 110 ESL participants after they read short texts from video trailers. The reading project that included 1) 10 video trailers (i.e., pre-reading activity), 2) 15 short reading passages with 94 target words that were recycled from the trailers, 3) hypertext-clickable words (each word included the part of speech, definition, example, and picture). The results show that repeated exposure to recycled vocabulary from the video trailers and short reading passages enhanced vocabulary gain and retention.

As for reading comprehension, Abraham (2008) carried out a meta-analysis of 11 studies on computer-mediated glosses in L2 reading comprehension and incidental vocabulary learning (see section 3.2.2 for incidental/intentional learning, p. 36). Data are collected from the results of the post-tests (immediate and delayed) and reading comprehension tests of all the studies in his meta-analysis. Cohen's d[1] is used to measure the effect sizes of these findings. The researcher reports a medium effect (between .21 and .80) on reading comprehension and a large effect (greater than .80) on incidental vocabulary learning. More specifically, Hegelheimer (1998), for instance, conducted a study examining the effect of textual glosses and sentence-level audio glosses on reading comprehension and vocabulary retention. The effect was tested through pre- and post-tests on a total of 115 ESL participants from different L1 backgrounds. The treatment included three online reading passages, where each passage is represented with different types of glosses: no glosses in the first reading

[1] Cohen's d is a standardized measure of effect sizes based on the differences between the means divided by the standard deviation.

passage, to textual glosses in the second reading passage, to textual and sentence-level audio glosses in the third passage. The findings show that the increased gloss mode did not yield any significant difference in reading comprehension. However, the use of two or more glosses resulted in significantly better performance in the vocabulary post-test.

More recently, Abraham (2007) investigated the effect of verbal and pictorial glosses on authentic text comprehension and new vocabulary learning. The participants (N=102; 61 females and 41 males) were divided into three groups: 1) control group (provided with the reading texts but no glosses), 2) choice look-up (given optional access to verbal and pictorial glosses for all the target words), and 3) forced look-up (forced to check glosses for all the target words). Participants were tested through pre-test, post-test, and recall protocols. The results revealed that the forced and choice look-up groups were significantly higher in the post-test and better in their recall protocols than the control group. However, no significant difference was found between the forced and choice look-up groups.

The previous literature review provides a general account of the use of glosses in developing L2 vocabulary and reading comprehension skills. Other research has been directed towards more specific topics such as look-up behaviour and global reading, word recognition skills, and form-meaning connections.

3.3.1.1 Look-up behaviour and global reading skills

Look-up behaviour includes research on learners' preferences for help option modes, online learning tools, and the visibility of help options. With respect to learners' preferences for help option modes, Lin and Chen (2007), for instance, argue for the dual-coding theory where multiple modes of help options are generally preferred by language learners because they activate two or more channels of perception (see section 3.2.4). The authors conducted an experiment in which they compared the impact of various types of visuals (i.e., animated versus static) and advance organizers (i.e., descriptive information or a question to activate prior knowledge) in an EFL private school in Taiwan. Study participants (N=115) were divided into four groups: 1) static visual only, 2) animation only, 3) animation plus descriptive advance organizer, and 4) animation plus question advance organizer. Generally, animation was found to have a

superior effect compared to the static condition. The animation plus question advance organizer is more effective than the other conditions for reading comprehension in both the post-test and delayed post-test.

The visibility of help options has also been investigated in a number of research studies and results generally indicate that visible links to help options are clicked more often than invisible links. Moreover, visible links are found to positively affect reading comprehension and speed. More specifically, De Ridder (2002), for instance, examined the influence of visibility/invisibility of electronic glosses (i.e., hyperlinks to dictionary definitions) in online texts on the flow of incidental vocabulary learning and reading comprehension. Sixty participants were divided into two groups: general subject matter reading sessions and specific (business topic) reading sessions. In each reading session, a text was presented in either visible (highlighted and underlined hyperlinks) glosses or invisible (unmarked) glosses. Testing included a vocabulary test, a comprehension test, and clicking and total time tracking. The study revealed that visible hyperlinks significantly increased the consulting of glosses without slowing down the reading speed. The increased amount of clicking did not affect incidental vocabulary learning or reading comprehension.

Similarly, Nikolova (2004) carried out a study comparing the effect of visible and invisible glosses in a computer reading environment on vocabulary acquisition and reading comprehension. The subjects (N=264) were grouped according to their language proficiency level (i.e., average or high). The texts were presented with either visible or invisible links to word meaning in the reading passage. Results indicate that visible links were more beneficial to average achievers than high achievers for both vocabulary acquisition and reading comprehension.

Another line of research is concerned with reading at the global level. The role of schemata, text structure and discourse organization, and extensive reading have been explored with various types of help options. For instance, Ariew and Ercetin (2004) compared word-level annotations to topic-level annotations with different forms of hypermedia. In their study, 84 advanced and intermediate ESL learners were asked to read a text for comprehension and to consult video or graphics annotations when

needed. Study results showed that the annotations did not affect reading comprehension for either level of learners. However, prior knowledge was found to positively affect reading comprehension. Similarly, Ercetin (2010) investigated prior knowledge and topic interest on text recall and annotations use. The annotations in the study included word-level and topic-level information. Data were collected through an immediate recall task, a topic interest questionnaire, a prior knowledge test, and semi-structured interviews. Study findings show that topic interest was more effective than prior knowledge for text recall. Also, a correlation was found between topic interest, prior knowledge, and topic-level annotations. The results showed that when the topic interest was low, learners with low prior knowledge tended to consult the annotations more. In contrast, when the topic interest was high, learners with high prior knowledge tended to consult the annotations more.

Another line of research has explored the use of definitions, syntactic information, anaphoric information, pronunciation, and supplements of reading context on discourse and text structure processing (e.g., Hulstijn, 2000; Kitajima, 2002; Van Gelderen, Schoonen, Stoel, De Glopper, & Hulstijn, 2007). This body of research emphasizes that annotation facilitates the processing of text in general and text structure in particular. Web 2.0 technology and dictionaries have also been investigated as online help options for extensive reading (Ducate & Lomicka, 2008; Maxim, 2002; Rankin, 2005). In these studies, qualitative data revealed that the learners were motivated to make their selection of readings and to read for pleasure outside class.

The following section reviews studies that examined word recognition skills through the use of help options in a computer reading environment.

3.3.1.2 Word recognition skills

Chun (2011) asserts that automatic recognition of L2 words and particularly high frequency words is a key component for successful L1 and L2 reading. Accordingly, Li (2010) investigated automaticity and word recognition in a CALL reading program with bilingual and monolingual dictionaries. Twenty Chinese ESL learners from grades 9 and 10 in a Canadian public school participated in the study. Their receptive vocabulary skills

were assessed with the vocabulary levels test by P. Nation (2001). Study results indicate that access to bilingual dictionaries enhances word recognition.

With respect to word recognition skills and working memory, Chun and Payne (2004) examined individual differences in working memory capacity and the use of multimedia annotations. Thirteen second-year L2 German students were asked to read a short story using reading software that provided multimedia annotations for difficult words. The learner's working memory capacity was assessed through a pseudoword repetition test and a reading span test. Their use of multimedia annotations was recorded through a tracking system, both while reading and doing vocabulary and reading exercises. Study results show that while learners with lower phonological working memory used multimedia annotations to compensate for their limitations, the difference between low- and high-reading span learners did not reveal significant results with regard to their reading comprehension.

Studies have also identified a beneficial role of help options for word recognition at the micro level and for enhancing decoding skills. Tozcu and Coady (2004), for instance, examined the relationship between explicit vocabulary learning and speed of word recognition. They conducted a study with 56 intermediate-level students studying English for the university preparatory program. The participants practiced on 2,000 high-frequency words using *New Lexis* software. Study results show that explicit learning in a CALL environment with help options decreases word recognition reaction time and, at the same time, improves reading comprehension and vocabulary acquisition.

The following section examines glossing with a focus on establishing a form-meaning connection for word learning.

3.3.1.3 Form-meaning connection

The common hypothesis is that if text glosses or annotations are consulted, the word is noticed first for its form and then for its meaning in the supplementary glosses (Peters, 2007). Form-meaning mapping requires documenting and testing of the cognitive processes involved in word reading, processing, and learning. According to VanPatten (2011), some researchers (e.g., Peters, 2007) confuse noticing with form-

meaning processing. VanPatten explains that learners may notice the word form, but this does not necessarily result in establishing form-meaning connections in the lexicon. The findings with regard to ESL Arabic learners, however, yield different speculations. Saigh and Schmitt (2012), Ryan and Meara (1991), Fender (2008), and Hayes-Harb (2006) suggest that Arabic learners do not process the vowel(s) in the word forms, which leads to confusion in establishing form-meaning connections. The meaning is always left unattached to a precise form; rather, it is attached to a group of word forms that have the same consonant strings.

Several studies investigated form-meaning connections in processing L2 words through glosses. For instance, Rott (2003) explored the role of multiple-choice glosses and multiple occurrences of a word as well as the role of the input-output cycle task on the form-meaning mappings for target words. After reading each short section in the passage, the participants, 14 L2 German learners, were asked to retell the content in German. The learners were first pre-tested on their knowledge of the target words and later post-tested on word gain. In addition, think-aloud protocols were employed to elicit L2 readers' processing behaviour. Results show that the gloss group differed significantly in gain and processing behaviour from the no-gloss group. For the reading behaviour, glosses were found to establish form-meaning connections; no glosses helped in global processing of the text but provided only shallow meaning mappings and skimming of words. Multiple occurrences of a word helped in refining and retaining the learners' knowledge of words and in establishing solid form-meaning connections for the target words.

In 2005, Rott investigated the role of multiple glosses in the quantity and quality of word processing. The concept of quantity and quality was first introduced by Hulstijn and Laufer (2001) with the involvement load hypothesis which predicts the depth of processing based on the amount of the involvement induced by the language learning task. Quantity referred to the number of short-term processed words and quality to those words that were deeply processed with form-meaning mappings. The participants were 10 English learners of German at a public university in the United States (Rott, 2005). They were in their third and fourth semesters of learning German. Target words were either enhanced with multiple glosses of information or with a single translation gloss in

English. Study results indicate that complete form-meaning connections were established by rehearsal of multiple information resources in the glosses or by semantic elaboration in the form of translation glosses. On the other hand, weak form-meaning connections were triggered by linear text processing where no additional sources of information were provided for the target words. Moreover, the multiple choice glosses helped the learners reach comprehension of even supporting ideas (at the global level).

In 1992, Hulstijn explored the role of glosses in lexical gain and retention. The main goal was to test the mental effort hypothesis, which states that retention occurs more when the meaning is inferred than when it is provided. Five experiments were conducted with two types of glosses: inferring meaning glosses and providing meaning glosses. Inferring meaning glosses were in the form of multiple choices for word meaning, concise context of the word meaning, or no cue/control cues for the word meaning. In contrast, providing meaning glosses were in the form of a word translation or a synonym. In all of the experiments, an unexpected post-test followed the participants' reading of the passage. The participants were Dutch learners from different L1 backgrounds. The general findings from his experiments assert that high mental effort (i.e., inferring meaning) is more effective in the robustness of form-meaning connections than is low mental effort (i.e., glossed translation). Uncontrolled reading with no cues, however, leads to incorrect inferences and incorrect establishment of form-meaning connections. Accordingly, Hulstijn concluded that glossing during reading helps in establishing and strengthening form-meaning connections as it induces high-mental effort in processing and hence, better retention.

In 1993, Hulstijn conducted a study to test word relevance and inferability, reading goals, and look-up behaviour. ESL Dutch students (N=82) were asked to read a text in a computer reading program with L1 translation glosses. The measurement tasks consisted of foreign language (FL) reading comprehension, inferring ability, and English vocabulary knowledge. Students' look-up behaviour was also analyzed via tracking data and video recording. The results showed that the learners' look-up behaviour was influenced by the relevance of the word, their ability to infer the meaning from the context, and their need for more information to fully understand the word. Based on these findings, retention is achieved if the word saturates the form-meaning connections'

requirements for lexical entry. According to Hulstijn, these connections are made when high effort is used to either infer or search for information in glosses.

Another hypothesis suggested by Hulstijn et al. (1996) relates to frequency of occurrence. The researchers designed a glossing study to test incidental acquisition of words occurring more than once in the text. They provided tools for checking the meaning of words. There were two experimental conditions (marginal glosses [L1 translation] and a bilingual dictionary) and one control condition. The researchers concluded that both frequency of occurrence and provision of meaning enhance retention of form-meaning connections. Hulstijn (2000, 2001) emphasized that hypertext/hypermedia glosses, whether in the form of an L1 translation or a bilingual dictionary, helped to establish form-meaning connections.

Recently, there have been a few attempts to implement either form-focused or meaning-focused glosses to enhance form-meaning connections of the target input. Sanko (2006), for instance, designed a CALL program with three reading conditions: a) meaning-focused annotations, b) form-focused annotations, and c) bilingual word lists. The study participants (N=120) were divided into 3 equal groups and given unknown target words, which were determined by a self-reporting test. The independent measures (receptive acquisition and retention) were tested by immediate and delayed (after three weeks) cloze tests which assessed the learners' knowledge of target words for both word meaning and form. The findings reported that the meaning-focused group outperformed the remaining groups, although the difference to the form-focused group was not significant.

In another attempt to emphasize form through glosses, Guidi (2009) studied the enhancement of form through frequency of the target form exemplars in text and L1 translation glosses. According to Guidi, noticing has never been addressed empirically in glossing studies. Thus, the major goal of his study was to investigate the role of glosses and type of linguistic items in noticing and learning. The experiment employed recognition and production in pre-, post-, and delayed tests. In addition, reading questionnaires and think-aloud protocols were utilized. The target forms included the present perfect and the impersonal *SE* in Spanish (i.e., a clitic to express impersonal

meaning). While the quantitative and qualitative measures of the study did not provide any evidence that L1 translation glossing stimulates noticing, the combination of glosses with other factors, such as salience of form, abstractness of meaning, and L1 and L2 grammaticalization patterns, was found to affect noticing and learning. In addition to glosses, Guidi used textual enhancement in his study to increase the salience of linguistic forms embedded in the input. The following section examines input enhancement and reviews existing research in this area.

3.3.2 Input enhancement

In addition to glossing, input enhancement is an important area in help option studies, and it is particularly relevant to this research study. The term input enhancement was first coined by Sharwood Smith (1993) and is defined as a set of implicit and unobtrusive techniques to increase the saliency of formal features in the L2 input (Jourdenais, Ota, Stauffer, Boyson, & Doughty, 1995).

Input enhancement is another form of reading assistance, which is commonly used for learning form rather than meaning. Drawing the learner's attention to the target language forms in a primarily contextualized environment is intended to help learners notice the gap between their interlanguage and the target language (Kim, 2003). It has been used as an attention-getting device in two forms: 1) typographical (textual enhancement) and 2) intonational (oral enhancement). Research has mainly focused on the first type (e.g., Leow, 1997, 2001; Shook, 1994; J. White, 1998). Input is generally typographically manipulated in three ways: 1) different fonts, 2) different sizes, and 3) different typographical effects, such as colors, bolding, italics, underlining, and capitalization. Sometimes, two or more techniques are used to create a stronger effect.

Input enhancement has been extensively studied in SLA, particularly in relation to grammar teaching and learning (Izumi, 2002; Jourdenais et al., 1995; Lee, 2007; Leow, 1997, 2001, 2007; Leow, Egi, Nuevo, & Tsai, 2003; Shook, 1994; J. White, 1998). Leow (2009a) and Han et al. (2008) provide comprehensive reviews of the input enhancement literature. Leow (2009a) divides research on input enhancement into two types based on their methodology: conflated enhancement research, in which the

study's independent variable is a combination of input enhancement and one or more other variables (e.g., feedback or explicit grammar instructions), and non-conflated enhancement research, in which the study compares the enhancement effect based on its presence or absence (+/-enhancement) with no conflation with other variables.

3.3.2.1 Conflated input enhancement research

According to Leow (2009a) and Han et al. (2008), this research combines input enhancement with various variables such as 1) recasts and oral or written interaction (Leeman, 2003; Sachs & Suh, 2007); 2) form-focused, meaning-focused, and/or processing instruction (Doughty, 1991; Leeman, Arteagoitia, Fridman, & Doughty, 1995; VanPatten, 1996); 3) corrective feedback (Lyddon, 2011; L. White, Spada, Lightbown, & Ranta, 1991); and 4) meaning comprehension and topic familiarity (Lee, 2007; Winke, 2013).

Recasts and textual input enhancement have been examined by both Leeman (2003) and Sachs and Suh (2007). Leeman (2003) explored the impact of enhanced recasts and positive evidence during oral interactions on the development of noun/adjective agreement markings of Spanish learners (N=74). The subjects were randomly assigned to four study groups: 1) recast, 2) negative evidence, 3) enhanced salience, and 4) control. Enhancement in the form of stress and intonation was present in the recast and the positive evidence group only. An oral-picture difference pre-test and immediate and delayed post-tests were administered to collect data on individual interactions between each participant and the researcher. The results indicated that the enhanced recast and positive evidence group were significantly better than the control group in both the immediate and delayed post-test. Sachs and Suh (2007) also investigated the effect of textual input enhancement of recasts during synchronous computer-mediated communication (CMC) between Korean EFL learners (N=30) and a native American English speaker. The EFL learners are divided into two groups: 1) enhancement group, and 2) no enhancement group. Enhancement was presented in the form of underlining and bolding the target form (shifting from simple past to present perfect). Data collection was through online think-aloud protocols, an offline multiple-choice comprehension test, and oral production through an interactive computer test. The findings reported no significant difference between the enhanced and unenhanced

group in target form accuracy. However, a high level of awareness of the target form was reported during the think-aloud protocols.

Input enhancement is also conflated with instructions on L2 grammar development, including form-focused, meaning-focused, and processing instruction (Doughty, 1991; Leeman et al., 1995; VanPatten & Cadierno, 1993a). For example, Doughty (1991) conducted a study in a computer-based environment on the role of instruction and enhancement in acquisition of the English relativization form. ESL learners (N=30) were split into three groups: 1) rule-oriented instruction, 2) meaning-oriented instruction, and 3) control group. In the rule-oriented instruction, the learners received grammatical explanations about relative-clause constructions in the study text. The meaning-oriented group were encouraged to pay attention to both content and structure (the target forms were highlighted and capitalized). The control group was provided with the reading text only. Using grammatical judgement and picture elicitation tasks, the study found that both instruction groups significantly improved more than the control group; however, the meaning-oriented group exhibited better comprehension of the reading text.

VanPatten and Cadierno (1993b) compared the effect of traditional rule-based instruction versus processing instruction. In both conditions, instructions were provided along with practice activities. However, the traditional instruction group received the feedback after production, while the processing instruction group received it during reading comprehension activities. The target structures were Spanish clitics and object pronouns. The results indicated that both instruction groups performed significantly better in the written production task even though the processing instruction group did not practice any production activities during the treatment. The processing instruction group also showed better comprehension than the other groups.

The effect of enhancement and instruction was also examined by Leeman et al. (1995), who compared the effect of form-focused instruction to communicative instruction. Form-focused instruction emphasizes meaning, but still draws attention to form through input enhancement techniques (e.g., colour highlighting and underlining). Communicative instruction makes no attempts to highlight form or to correct language

errors. The target form was preterite versus imperfect in Spanish, which was tested through debate, essays, and a modified cloze test. The form-focused group exhibited significant improvement in target form accuracy. The communicative group only showed improvement in the essay task.

Corrective feedback has also been studied in conflation with input enhancement (Lyddon, 2011; L. White et al., 1991). L. White et al. (1991) examined accuracy in question formation in 129 francophone beginner ESL students using form-focused instruction and corrective feedback with input enhancement. The treatment group received 3 hours of form-focused instruction followed by 2 hours of input enhancement activities and corrective feedback. The control group had the same reading materials but were not instructed on form. The treatment group outperformed the control group in accuracy of question formation. L. White et al. (1991) stressed that input enhancement can trigger significant changes in the learner's interlanguage system. Similarly, Lyddon (2011) investigated the learning of the *à/au/en/aux* distinction in an interactive computer-based reading environment. The reading lessons provided different types of textually enhanced or unenhanced corrective feedback. The French learners (N=136) were divided into four treatment groups: "meaning-focused, implicit form-focused, non-metalinguistic explicit form-focused, or metalinguistic explicit form-focused" (p. 104). The study reported significant improvement for all types of feedback with and without textual input enhancement, suggesting that all techniques can be equally beneficial.

Reading comprehension and topic familiarity have been examined in several input enhancement studies. Lee (2007), for instance, attempted to deliver meaning-focused reading classes, in which English passive forms were brought to the learners' attention without interrupting the meaning comprehension process. The subjects, Korean ESL learners (N=259), were assigned to four treatment groups basd on the presence or absence of textual input enhancement or topic familarity conditions. The study focused on the accuracy of identifying and correcting English passive forms and on the level of reading comprehension. Textual input enhancement resulted in the accurate use of the target form, but it led to a trade-off in meaning comprehension. On the other hand, topic familarity improved reading comprehension but not the learning of the target form.

Winke (2013) replicated Lee's (2007) study with two modifications. This study used an independent reading proficiency test and eye-tracking data for noticing and processing of the target form. Intermediate ESL learners (N=55) were enrolled in a reading session of an authentic text flooded with the target form (i.e., English passives). The findings revealed no significant difference between the study groups in acquistion of the target form or reading comprehension. However, the eye-tracking data revealed a significant effect of input enhancement on the learners' total fixation duration as well as on re-reading times.

Notably, the findings in the conflated research are inconsistent. Leow (2009a) explained these inconsistent findings based on two factors. Firstly, he argued that the conflation of one or two variables combined with input enhancement confounds the impact of enhancement alone. Secondly, the lack of the concurrent data in conflated research makes it difficult to determine the extent to which input enhancement is effective when conflated with other variables. Only two studies, Sachs and Suh (2007) and Winke (2013) provide an empirical account of the role of enhancement on noticing and learning the target form. Both studies reported the efficacy of input enhancement on noticing the target forms, an effect that was not, however, reflected in the quantitative data.

3.3.2.2 Non-conflated input enhancement research

In non-conflated studies, the research design is based on the comparison between enhanced and unenhanced forms, generally allowing a better determination of the effect of input enhancement. Differences in the research methodologies are based on "quite a range of linguistic items, languages and language levels, amount of exposure, text lengths, and postexposure linguistic assessment tasks" (Leow, 2009a, p. 23).

Jourdenais et al. (1995) investigated the impact of enhanced input on the written production of Spanish preterite and imperfect verb forms. All the instances of the Spanish preterite and imperfect forms were highlighted in the treatment group; the control group had no enhancement. Adult English learners of Spanish (N=10) in their second semester participated in the study. The post-experimental task was a picture-

based writing task and a reading-comprehension task, followed immediately by offline think-aloud protocols. The qualitative data from the oral protocols showed that the enhanced group used the past tense more than the control group. However, there was no evidence of the enhancement effect of target aspectual alternations on the functionality between the forms. Analysis of the elicited written data confirmed the qualitative findings: the past tense was used more in the production of the enhanced group than in that of the control group. Enhancement did not have any effect on the use of the target forms as distant aspectual markers. The researchers concluded that enhanced input showed a priming effect, facilitating noticing and learning the past tense as a subsequent part of the target form, and accordingly, enhancement is found to be a positive facilitator of L2 form learning. However, Jourdenais (1998) carried out the same study in a larger population (N=124) that did not show a significant effect of input enhancement.

These conflicting results are confirmed by Overstreet's (1998) study on L2 Spanish learners (N=50) targeting the preterite and imperfect verb forms. Enhancement was provided in the form of underlining, bolding, shadowing, and contrasting font. Using an offline oral picture description task, the enhancement group did not show any positive effect in either production or comprehension.

Leow (1997) considered text length in his study on the impact of input enhancement on intake during a reading comprehension task. The subjects were 84 English learners of Spanish in their second semester of language learning. The target form was the Spanish formal imperative, tested through a multiple-choice recognition task and a short-answer comprehension task. The length of the passage was found to have an effect on reading comprehension; however, no significant effect was observed for input enhancement. Measures of intake did not show any significant effect for text length and input enhancement.

Alanen (1995) studied the effect of exposure to explicit and implicit input both with and without visual enhancement (i.e., use of italics). The target forms were three Finnish locative suffixes and four types of consonant alternations presented in an online reading comprehension task. The participants were 36 English learners of Finnish at the

University of Hawaii, Manoa. Data collection consisted of online and offline measures of think-aloud protocols and sentence-completion tasks. Data analysis revealed that typographical enhancement of the input did not have an effect on the learning of target forms, especially in the case of implicit learning. The production data did not show any evidence of form-function mappings. The explicit condition was found to be effective for learning in all measures and in both the enhanced and control groups. Hulstijn (1992) asserts that incidental learning requires more mental effort and time, an assertion that explains Alanen's results. Hence, post-experimental measures should allow time for testing production and enhanced input. J. White (1998) reported similar results on the impact of enhancing English possessive determiners in reading passages. French ESL learners (N=86) participated in the study for a period of two weeks (10 hours in total). The tasks were multiple choice, text correction, and oral picture-description. Enhancement led to a partial effect in acquisition; however, the results were not statistically significant.

Another factor in input enhancement is type of linguistic form. A limited number of studies have investigated the type of linguistic form as an independent variable. For instance, VanPatten (1990) examined the effect of regular and simplified input on the intake of different linguistic items (present perfect and subjunctive) at two different levels of language learning (first and fourth semesters). The participants were college students majoring in Spanish. Intake was measured by a multiple-choice recognition task. The target linguistic forms were chosen based on the hypothesis that more salient phonological and meaningful forms will be attended to before other forms. The findings revealed no evidence to support this hypothesis, as more experienced learners showed more intake of forms in general. Guidi (2009) argues that the morphological nature of both form endings drew the learners' attention to the lexical meaning of the verb rather than the functional connections of form. However, in one of his early studies, before formulating his input hypothesis, VanPatten (1990) examined the processing of form and content of Spanish learners listening to audio clips. The learners were given explicit instructions to process information under four conditions: a) attending to content, b) attending to meaning and important vocabulary items, c) attending to meaning and syntactic information (i.e., definite articles), and d) attending to meaning and morphological information (i.e., verb forms). The dependent measure was a recall

protocol written in the L1 (English). Study results revealed that attending to form and content in the early stages of language learning was difficult in processing. Moreover, morphological information of the verb endings was found to have the most negative effect on input processing. VanPatten concluded that morphology lacks communicative value, making it the part of language least noticed and perceived during normal L2 listening settings.

There is a consistent body of findings in the non-conflated input enhancement research suggesting that there is no effect from enhancement when compared to the control group. The only two studies that produced a significant effect are by Shook (1994) and Alsadoon and Heift (2015). Shook (1994) based his study on the communicative value and meaningfulness of form. He investigated the recognition of form when learners were exposed to varying degrees of *meaningfulness* in form. The target forms were Spanish present perfect (partially meaningful with tense and aspect) and Spanish pronouns (less meaningful). The target structures were presented bolded and capitalized in one of two reading passages. The study also included a rule-formulation task while reading the enhanced text. Study results revealed that more explicit enhancement along with the rule-production task were beneficial for noticing and processing of *meaningless* forms. Leow (2009a) argued that this study produced different results due to the effect of the rule-production task, which triggered higher noticing and deeper processing of the target forms.

A very closely related study to this dissertation is by Alsadoon and Heift (2015) who attempted to reduce vowel blindness in Arabic ESL learners by providing textual input enhancement in the form of highlighting of the vowels. The study mainly examined the impact of input enhancement on the noticing and intake of English vowels by Arabic ESL learners. The study employed eye-tracking to record the learner's eye fixations on vowels during the reading of short sentences. Thirty beginning Arabic ESL learners were divided into an experimental group with vowel enhancement and a control group with no enhancement. The study reported a significant effect of using textual input enhancement on reducing vowel blindness of Arabic ESL learners. Moreover, it found that textual input enhancement attracted learners' attention with longer fixations on the target words. The researchers interpret these significant results based on methodological differences

between the study design and previous studies in the non-conflated strand including the use of short sentences, familiar target words, and the type of the target linguistic form.

Leow (2001, 2007) has argued that there is a methodological design problem in the internal validity of most measures used in input enhancement studies. The main issue with the internal validity is that attention processes are not set to properly measure enough empirical data during experimentation. He argued that concurrent online think-aloud protocols should be employed to elicit data about the processing of form rather than the production of form. Therefore, Leow (2001) designed a study in which online think-aloud protocols were used to report noticing of the enhanced input versus the unenhanced condition. The elicitation data was compared with intake from a multiple-choice recognition task. The target form was the Spanish formal and polite imperatives. The subjects consisted of 74 English beginner learners of Spanish. Findings from the think-aloud protocols showed that noticing was the same for both group conditions (enhanced versus unenhanced). However, there was positive intake by the enhanced group. This study was the first to distinguish between the role of external enhancement on the internal mechanism of noticing (Guidi, 2009). Bowles (2003) replicated Leow's study and reported the same lack of effect for textual input enhancement. Leow (2009a) stressed that think-aloud data provided an explanation of the consistently insignificant results in non-conflated input enhancement research. Using think-aloud protocols, both the studies of Leow (2001) and Bowles (2003) found through concurrent think-aloud data a statistically equal amount of noticing to occur in both enhanced and unenhanced conditions. They explained that the learners may read and notice the target form for meaning rather than for form, resulting in a low level of awareness processing of the form.

Izumi (2002) used note-taking to record concurrent data from 61 ESL participants while reading five passages over a two-week period. The study participants were encouraged to write down notes of words that helped them in text comprehension. The target forms were English relative clauses, which were either enhanced (experimental group) or non-enhanced (control group). The study reported no significant results, though enhancement prompted noticing. Winke (2013) claimed that the length of relative

clauses affects their noticeability, making it hard to determine whether form length or enhancement induced the learners' noticing.

In an attempt to employ more accurate concurrent elicitation methods, Winke (2013) used eye-tracking technology to record the participants' eye-movements and fixations (See section 3.3.2.1). The results showed that textual input enhancement attracted longer duration fixations and re-reading durations; however, this enhanced noticing was not reflected in the learners' gain scores. Alsadoon and Heift (2015) confirm Winke's finding but they report further that textual input enhancement is effective for the learners' deep processing and intake of the target forms.

3.4 Summary

This chapter examined the literature in help options and CALL, mainly, in the areas of vocabulary learning and reading comprehension. First, it reviewed the theories that are most commonly used to inform help option research in SLA and CALL. Then, it covered research in glossing and input enhancement, as they are most relevant to the help options used in this dissertation study. The glossing section discussed vocabulary learning and reading comprehension studies with respect to learners' look-up behaviour and global reading, word recognition skills, and form-meaning connections. Finally, input enhancement reviewed the studies based on two research strands: conflated and non-conflated input enhancement research.

The study of this dissertation uses a non-conflated condition of input enhancement in the form of highlighting and underlining. It aims to contribute to the input enhancement literature particularly with respect to vowel blindness of Arabic EFL learners. The following section identifies research gaps in the existing literature on help options and outlines the contributions of the current study to the fields of CALL and SLA.

3.5 Research gaps

The literature review on vowel blindness presented in chapter 2 identifies vowel blindness as a fairly young research area with limited empirical research. Therefore, the vowel blindness phenomenon, in general, needs to be studied extensively and from different angles and with different perspectives. More specifically, evidence about the existence and nature of vowel blindness has encouraged training studies to help Arabic ESL learners to overcome the problem (Saigh & Schmitt, 2012; Stein, 2010; Taylor, 2008). The study by Taylor (2008) is the only training study which provides computer-based reading phonics and spelling. The current study contributes to the vowel blindness area in two ways: a) adding more insights about the nature of the problem and b) experimenting with different help options to address and overcome the vowel blindness problem.

The literature review on help options provided in chapter 3 indicates that, generally speaking, help options are found to have a positive impact on reading comprehension and vocabulary learning (Chun, 2011). However, there are several gaps in the glossing and input enhancement literature with respect to form and meaning.

Gaps were identified in form-meaning connection research. Form-meaning studies are scarce, and only two researchers (Hulstijn, 1992, 1993; Rott, 2003, 2005) have explored its relevance to L2 Learning. While the findings generally reveal that glossing helps in establishing a form-meaning connection, the mechanisms that are most beneficial for long-term retention need to be further explored. Moreover, in these studies, the CALL environment was not especially designed and targeted to focus on the form-meaning connections. Instead, the research focused on other variables, such as incidental learning, frequency of occurrence, and load involvement. Therefore, attention must be paid to the design of the glossing environment in order to specifically study its impact on establishing form-meaning connections.

Some researchers (e.g., Guidi, 2009; Sanko, 2006) examined information-focused glosses by developing CALL reading programs that provided specific information for either form or meaning. Their findings support the notion that information focused on form or meaning targets deficiencies in word-meaning connections.

Additionally, it promotes noticing and processing of new words. However, form-focused glosses have not been explored with respect to target form problems such as those related to vowel blindness, grammatical forms, or morphological structures. Yet, these glosses may be an effective tool in contextualizing form information, prompting noticing, and compensating for form-related deficiencies in L2 decoding and encoding skills.

On the other hand, there is a wealth of research that relates to input enhancement, particularly in grammar (Alanen, 1995; Leow, 2007; Shook, 1994). These findings emphasize that noticing of target forms is more likely to occur in an enhanced environment (Jourdenais et al., 1995; Leow, 2001, 2007; Leow et al., 2003; Shook, 1994). However, there is only one study that investigated the impact of textual input enhancement on vowel blindness, that of Alsadoon and Heift (2015). The study employed a small number of participants (N=30) in an ESL context. The literature needs more empirical evidence with respect to input enhancement for reading and/or spelling, and with respect to noticing and processing of fine-grained forms such as phonemes (e.g., vowels).

Accordingly, the current research attempts to fill in the above-mentioned gaps in the literature. The overall goal is to address the vowel blindness phenomenon of Arabic learners of English through help options, namely form-focused glosses and input enhancement for vowels. The glosses consist of different types of phonological grapheme-phoneme and grapheme-syllable information, while input enhancement highlights the vowels in the target words. Consequently, it seeks to reduce vowel blindness by promoting noticing of the target form and establishing form-meaning connections. As for the size of the study population, it studies a larger sample of participants than previous research (e.g., Alsadoon & Heift, 2015) and tests these help options in an EFL context by using a mixed-method research design.

4 Research methodology

The research conducted for this dissertation addresses the problems of vowel blindness for Arabic EFL learners described in chapter 2 coupled with the gaps in help options identified in chapter 3. This chapter presents the research questions and provides a description of the study participants. It then describes VALE, the CALL program designed by the author of this dissertation, which guides learners through all the study instruments and stages. VALE both delivered most of the data gathering instruments of the study and recorded participants' responses, and implemented the control treatment and the four experimental treatments designed to assist Arabic ESL/EFL learners in improving their English vowel reading. The study methodology for data collection and analysis is also described.

4.1 Research questions

The main goal of the study is to investigate the potentially different effects of four distinct help options on vowel blindness in a CALL reading environment. The core of VALE provides two broad types of help options, namely typographical input enhancement and form-focused glosses. Typographical input enhancement is achieved by highlighting the vowel(s). Form-focused glosses are of three types: 1) segment focused glosses; 2) syllable focused glosses; and 3) segment-syllable focused glosses. The research for this dissertation addresses the following research questions which are grouped into three different sets.

The first set of research questions investigates the effect on vowel blindness of VALE's help options for targeted words (i.e., those directly treated with the help). More specifically:

1.1: Over the initial intervention period (pre-test to post-test), do four different help options (in the form of input enhancement and three types of form-focused glosses) reduce vowel blindness more than no help at all? Do the four options differ in their effects on vowel blindness reduction?

1.2: Over the retention period (post-test to delayed post-test), do the four different help options lead to maintenance of any reduction in vowel blindness better than no help at all? Do the four options differ in their effects on maintenance of vowel blindness reduction?

1.3: Over the entire period of the study (pre-test to delayed post-test), do the four different help options yield a net reduction in vowel blindness more than no help at all? Do the four options differ in their effects on overall vowel blindness reduction?

The second set of research questions examines the effect on vowel blindness of VALE's help options for untreated words. More specifically:

2.1: Over the initial intervention period (pre-test to post-test), do four different help options (in the form of input enhancement and three types of form-focused glosses) reduce vowel blindness of nontarget/untreated items more than no help at all? Do the four options differ in their effects on vowel blindness reduction of nontarget/untreated items?

2.2: Over the retention period (post-test to delayed post-test), do the four different help options lead to maintenance of any reduction in vowel blindness of nontarget/untreated items better than no help at all? Do the four options differ in their effects on maintenance of vowel blindness reduction of nontarget/untreated items?

2.3: Over the entire period of the study (pre-test to delayed post-test), do the four different help options yield a net reduction in vowel blindness of nontarget/untreated items more than no help at all? Do the four options differ in their effects on overall vowel blindness reduction of nontarget/untreated items?

Finally, the third set of research questions investigates the study participants' awareness of vowel blindness and attitudes towards VALE. More specifically:

3.1: How satisfied were the learners with the technical design features of VALE?

3.2: How far did the learners perceive the help options of VALE as raising their awareness of vowel blindness problems?

3.3: How far did the learners perceive the help options of VALE as assisting their learning about vowels?

4.2 Participants

Two hundred-fifty Saudi Arabian EFL students at a beginner to low-intermediate level participated in the study. The participants were recruited from the English program in the foundation year at Al-Imam Mohammed Bin Saud University in Riyadh, Saudi Arabia. The students' proficiency level was determined prior to the study by the program's placement test[2]. The university's English program lasts for two semesters over a total of 8 months. The first semester is aimed at beginner to intermediate EFL students and the second level is for intermediate to advanced EFL learners.

The participants in the current study were all registered in the first semester of the program. They were recruited through email and in-class invitations by the researcher and by their instructors. Initially, the sample was 261 participants but four were eliminated from the study because they indicated a high-intermediate level of proficiency in the background questionnaire through having taken additional private intensive EFL classes. A further seven participants exercised their right to withdraw from the study at different points: 4 after taking the pre-test and 3 after the post-test.

The participants were recruited in the seventh week of their program for the following two reasons. Firstly, this ensured that the participants had settled into their academic program and environment, and secondly, this allowed sufficient exposure to the new type of EFL instruction they were receiving. Most of the students had previously received EFL instruction for six years during their secondary school education which predominantly focuses on the basic alphabet, grammar, and vocabulary of EFL. The foundation year EFL program, however, is a transitional program which, unlike their basic EFL training, exposes learners to the more systematic teaching of all language skills including reading, speaking, listening, writing, as well as grammar. It is meant to fill

[2] The placement test examined the English grammar, vocabulary, reading, and listening skills of the applicants. It includes 100 items that are either fill in the blanks or multiple-choice questions. One score is given for each test item, scores from 20-40 indicate beginners to low-intermediates and those students are enrolled to the intensive English program. However, scores higher than 40 indicate intermediate level or higher level and those students are enrolled to level 1 where they can take some other academic courses in addition to the EFL courses.

the gap between the (often little) English they know on leaving high school and the English they need in order to pursue their majors through the medium of English in the following year since increasingly Saudi Arabian universities attempt to deliver many of their BA programs in English.

The study participants were all females as the educational system in Saudi Arabia segregates females and males so direct access by a female researcher to male participants is not possible. Cultural norms were respected in the design of the research instruments, such as avoiding video recordings and/or taking pictures.

As part of the data collection, study participants were asked to fill in a background questionnaire which collected detailed information about their age, English background, computer skills, and their experience with CALL software. Table 4.1 provides detailed information on the study participants' background (see Appendix A for the background questionnaire). The participants' ages ranged from 18 to 21 years with an average of 18.6 years old. Most of the participants' previous exposure to English other than in English classes in Saudi state schools did not exceed 0-2 months (83.2%), with only 19 participants (7.6%) having studied abroad for a short period of time, for a maximum of four months. As for other languages, only 8 participants (3.2%) claimed to know a third language. These languages were French (2%), Turkish (0.8%), and Spanish (0.4%).

With respect to their CALL experience prior to the study, the majority of the participants (81%) claimed to use the Internet particularly for English reading and vocabulary. However, English books were still found to be used at almost the same level (78%). Social networking was also moderately used by the participants (55%) in improving English reading and vocabulary. Tutors, friends, and computer programs seem to range at the same lower level of usage as sources of additional English exposure/instruction (from 32% to 18%).

Table 4.1. Background information on the study participants

Characteristics	Range and frequency	Percentage of sample
General information		
Number of participants	250	
Age		
Range	18-21 years	
Mean	18.6 years	
Gender	All females	
Academic status	All freshmen	
English background		
Duration of Studying in ESL/ EFL schools (not including secondary education)		
9-11 months	1 participant	0.4%
6-8 months	9 participants	3.5%
3-5 months	32 participants	12.8%
0-2 months	208 participants	83.2%
Studying English abroad		
Yes	19 participants	7.6%
No	231 participants	92.4%
Duration of studying abroad		
0 months	231 participants	92.4%
3 months	6 participants	2.4%
4 months	3 participants	1.2%
2 months	10 participants	4.0%
Other languages known		
Any other languages?		
Yes	8 participants	3.2%
No	242 participants	96.8%
Which language?		
French	5 participants	2.0%
Turkish	2 participants	0.8%
Spanish	1 participant	0.4%

Characteristics	Range and frequency	Percentage of sample
Prior CALL experience		
Resources used for English reading practice		
English books	197 participants	78.8%
Internet	203 participants	81.2%
Computer programs	77 participants	30.8%
Social networking	138 participants	55.2%
Tutor	61 participants	24.4%
Friend	60 participants	24.0%
Others (no examples are provided by the learners)	18 participants	7.2%
Resources used for English vocabulary practice		
English books	195 participants	78.0%
Internet	204 participants	81.6%
Computer programs	96 participants	38.4%
Social network	121 participants	48.4%
Tutor	45 participants	18.0%
Friend	81 participants	32.4%
Others (e.g., cell phone apps, video games, movies, and puzzles).	62 participants	24.8%
Which do you access more often?		
Print-based materials	58 participants	23.2%
Computer-based materials	192 participants	76.8%
Which do you prefer?		
Print-based materials	89 participants	35.6%
Computer-based materials	161 participants	64.4%
Do you use the internet or computer tools for reading such as multimedia, subtitles/ transcripts, reading exercises, translation, and other reading tools?		
Yes	222 participants	88.8%
No	28 participants	11.2%
Are they useful?		
Very useful	83 participants	33.2%
Useful	119 participants	47.6%
Somewhat	13 participants	5.2%
Not particularly useful	10 participants	4.0%
N/A	25 participants	10.0%

Characteristics	Range and frequency	Percentage of sample
Do you use the Internet or computer tools for learning vocabulary such as dictionaries, word pronunciation, translation, and other online tools?		
Yes	235 participants	94.0%
No	15 participants	6.0%
Are they useful?		
Very useful	102 participants	40.8%
Useful	120 participants	48.0%
Somewhat	10 participants	4.0%
Not particularly useful	4 participants	1.6%
N/A	14 participants	5.6%

4.3 Study design

In order to measure the effects of input enhancement and form-focused glosses on vowel blindness, a mixed-method research approach was implemented which gathers quantitative as well as qualitative data to answer the research questions. The study further employs a between-subjects experimental design with four treatment groups and one control group. Group assignment was random and based on the two main types of help options implemented in VALE: 1) input enhancement and 2) form-focused glosses. Form-focused glosses further divide into three sub-groups: segment-focused glosses, syllable-focused glosses, and segment-syllable focused glosses.

The key independent variable in this study is the type of intervention received by a participant while silently reading during the treatment phase. The treatment variable had five values: input enhancement, or one of three types of form-focused glosses which participants were required to access, or none of those. The input enhancement variable provided participants with four reading passages in English in which vowels in the target words were typographically altered to enhance vowel salience. The form-focused glosses, on the other hand, provided participants with segment, syllable, or segment-syllable vowel information for the target words in the same four reading passages.

The dependent variable in this study is vowel blindness of Arabic EFL learners which was pre-tested with a form-meaning multiple choice recognition task for 32 target words and 8 nontarget words (i.e., distractors) to confirm the learners' problem with vowel blindness. After the treatment phase, the study participants took an immediate post-test and 6 weeks later a delayed post-test. The same test was used with distractors for an analysis of untreated words. The study also attempted to capture the learners' learning experience and their attitudes towards VALE through a questionnaire and retrospective interviews that were conducted with 40 participants at the end of the experiment.

Table 4.2. Summary of the research design

<table>
<tr><td rowspan="2">Pre-treatment</td><td colspan="9">Main treatment and 5 treatment groups</td></tr>
<tr><td colspan="9">Between-subjects factors</td></tr>
<tr><td rowspan="4">Consent-form
+
Background questionnaire
+
Tutorial</td><td colspan="5">Dependent Variable</td><td colspan="4">Values of treatment variable</td></tr>
<tr><td colspan="5" rowspan="2">Vowel blindness</td><td>Input enhancement</td><td colspan="2">Form-focused glosses</td><td colspan="2">No help options</td></tr>
<tr><td>Highlighting vowels</td><td>Segment</td><td>Syllable</td><td>Segment -syllable</td><td>Control group</td></tr>
<tr><td colspan="4">Quantitative measures</td><td>Qualitative measures</td><td colspan="4" rowspan="2">Reading task

4 sessions (4 reading passages + word-meaning exercise)</td></tr>
<tr><td>Pre-test</td><td>Post -test</td><td>Delayed post-test</td><td>Attitude Quest-ionnaire</td><td>Retro-spective interviews</td></tr>
<tr><td>30 mins</td><td>30 mins</td><td>30 mins</td><td>30 mins</td><td>30 |mins</td><td>15 mins</td><td colspan="4">20 mins for each session making a total of 80 mins for all four sessions</td></tr>
<tr><td>Total time</td><td colspan="9">245 mins</td></tr>
</table>

Table 4.2 illustrates the research design and sequence of procedures of the study. It depicts the between-subjects and mixed-method research design and specifies the time allotted for the entire study (245 minutes) as well as for each individual session.

The study consists of three phases: pre-treatment (30 minutes), treatment (total of 140 minutes), and post-treatment (total of 75 minutes). The entire study took place over three days. Day 1 was used for the pre-treatment, day 2 for the main treatment, and day 3 for the post-treatment which took place after 6 weeks.

The three study phases of pre-treatment, treatment, and post-treatment were delivered to the study participants largely through VALE. The following sections describe the design and functionality of VALE including the program flow, tests, questionnaires, reading passages, target words, and help options.

4.4 VALE

The pre-treatment phase included a video tutorial on how to use VALE, an electronic consent form, and a background questionnaire. The treatment phase was the longest phase in the study and it included a pre-test, four reading sessions, and a post-test. Finally, the post-treatment, which took place after six weeks, entailed a delayed post-test, an attitude questionnaire, and (outside VALE) retrospective interviews. The learners accessed VALE through Al-imam Mohammed Bin Saud University's computer labs during their regular reading classes' hours.

4.4.1 Pre-treatment phase

The participants were first asked to log on to the program's website at http://www.v-a-l-e.com. Figure 4.1 displays the welcome page which gives general information about the study and a link to a tutorial video on how to register in and use the program. There is a drop-down sign-up interface for new users, or log-on fields for username and password for returning users. From this page, the learners first click on the video tutorial to obtain step-by-step instructions on how to use VALE which was followed by a face-to-face open session between the researcher and the participants to answer questions about VALE and the study in general. In addition, the study participants were given detailed information and instructions about VALE and the study during the recruitment sessions.

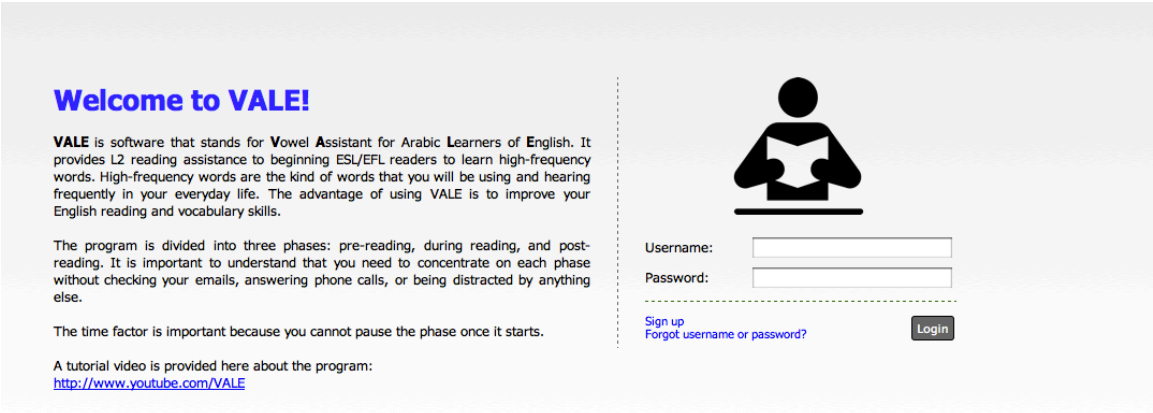


Figure 4.1. VALE welcome page

After registering and logging on to VALE, the main menu page (see Figure 4.2) provided an overview of the study with its three phases. The main menu page is divided into four columns and three rows. The columns give information about the day, the time, and the status of the study phase (to do, completed, and to start next). The rows show the components of each study phase. For instance, phase 1 consists of a consent form and a questionnaire. Immediate tasks are highlighted in light green (phase 2) while either completed (phase 1) or future tasks (phase 3) are highlighted in grey.

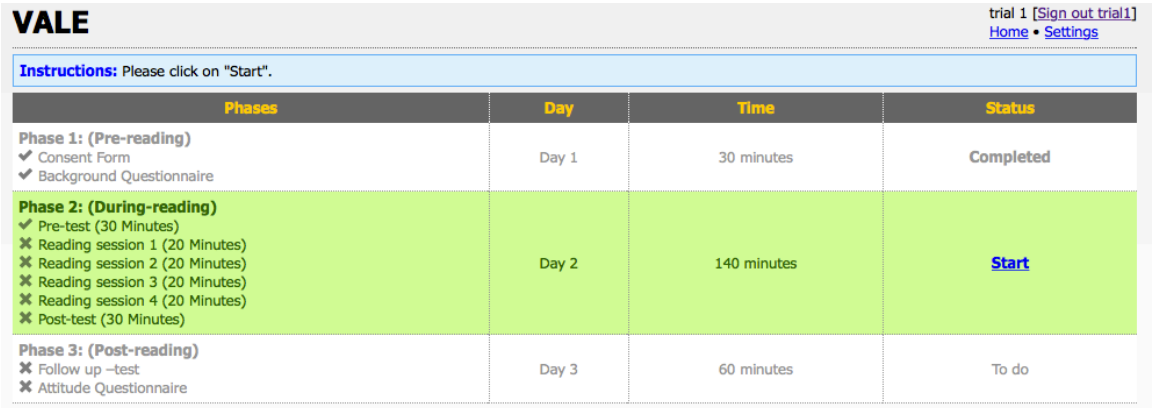


Figure 4.2. Main menu

Figure 4.2 further shows that the pre-treatment phase lasted 30 minutes. The pre-treatment phase consisted of an online consent form which explained the nature of the study as well as the participants' right to withdraw at any time during the study. Study participants were also informed that their information would be kept confidential and

would not be shared with their teachers thus possibly affecting their course grades. After signing the consent form, study participants watched a 7-minute step-by-step video tutorial on how to use VALE. They were informed that they were free to watch the tutorial video again at anytime. After this, a background questionnaire elicited demographic information on the study participants including their English language background, their general computer use as well as their use of online help options for language learning such as dictionaries, translation tools, etc. (see Table 4.1 and Appendix A).

4.4.2 Treatment phase

The treatment phase started on day 2 with a pre-test of both the targeted and nontargeted words. Study participants were given 30 minutes to complete the test. After finishing the pre-test (see section 4.4.2.2), the reading session started which consisted of four reading passages, each with an exercise to complete which checked knowledge of, or ability to, guess a target word's meaning.

The reading passages were provided in one of the four experimental conditions or with no help other than the word meaning for the control group. For instance, Figure 4.3 displays the second reading passage, *True Friends*, for the experimental group that received segment-focused glosses for the targeted words. These glosses are displayed below the reading passage when a word is clicked. For each reading passage, study participants were given 20 minutes to complete the task. A timer indicating the remaining time was displayed in the upper right hand corner of the display window. To ensure that the study participants paid attention to and checked the glosses of the target words, they were not able to advance to the next reading passage until they had accessed all glosses. This procedure was applied to all the reading passages. Furthermore, and to encourage learners to read the texts for meaning and not just pay attention to individual word forms, in all five treatment conditions a multiple choice exercise that tested word meaning was provided in the right half of the display window. If the learner provided an incorrect answer on word meaning, the program forced her to check the meaning from the meaning options provided in all conditions before moving on to the next reading passage.

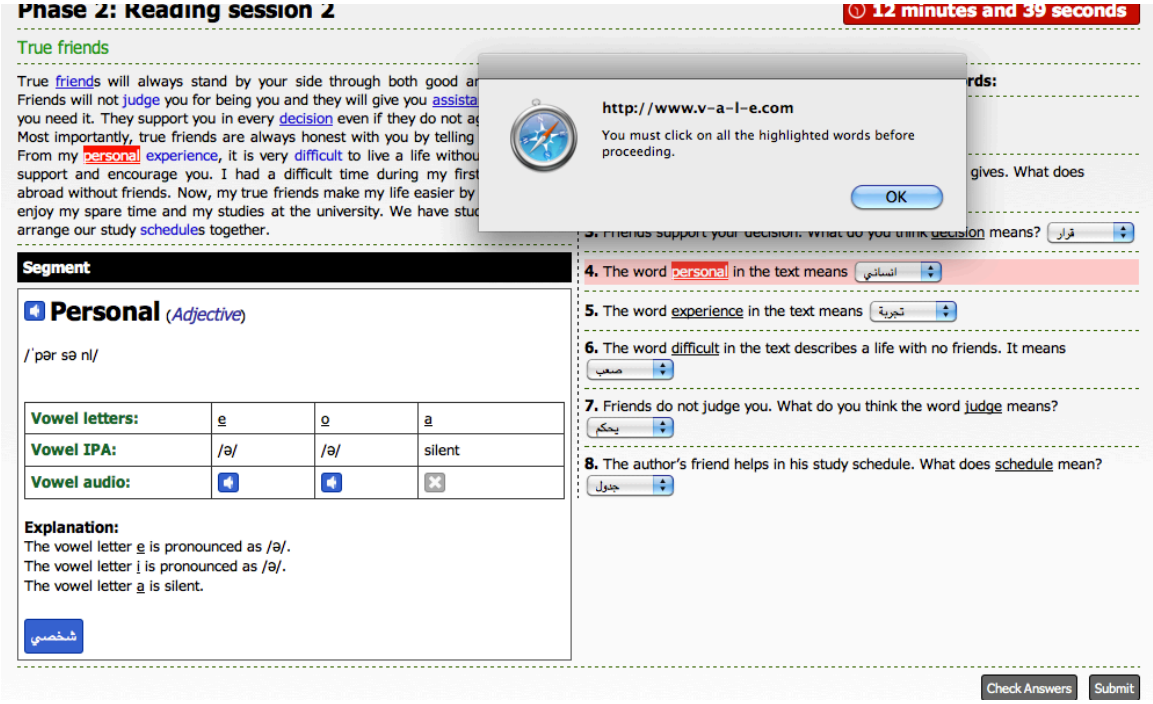


Figure 4.3. A reading passage session featuring the word meaning m/c exercise with a system alert to click on word meaning in the gloss for an incorrect answer

4.4.2.1 Reading passages

The four reading passages were written by the author of this dissertation with consultation from an expert editor, an ESL teacher, and two graduate students from the Linguistics Department at Simon Fraser University.

Each text contains 9–10 lines ranging from 113 to 130 words. The total number of words in all four texts adds up to 490 words which, according to Horst (2005)[3], is within the range of participants with a beginning EFL reading level. Four texts instead of one text were chosen to spread out the 32 target words and to not unnecessarily overwhelm the readers with 32 words and their glosses all at the same time. According to Hulstijn (1993) and Hulstijn and Laufer (2001), the lighter the general processing load for reading, the more attention is given to word-level reading skills.

[3] Horst (2005) states that texts of 400–600 words are appropriate for beginning level readers.

As for the choice of reading topics, Davis (1989) stresses the importance of topic familiarity when choosing texts for ESL readers, especially for beginning students. He emphasizes that widely known background information and culturally familiar topics should be utilized for beginner readers in order to promote the learners' reading involvement and to lighten the text level reading burden. Accordingly, the current study adopted culturally and otherwise universally familiar topics such as traditional clothing in Saudi Arabia, friendship, sport, and Arabian coffee (see Appendix C).

Finally, syntactic ease was another key factor in the construction of the texts. For instance, syntactically simple structures, or explicit pronouns, and explicit pragmatic reference-relations were used. Therefore, the goal was to make the processing load for the working memory lighter with short texts, familiar topics, and easy syntax. Hence, it is expected to leave more space available in the working memory for dealing with meaning of unknown words or the form of known words. Table 4.3 provides an overview of the four reading texts and their specifics.

Table 4.3. Summary of the study reading passages

Reading	Topic	Genre	Length	Tense
Text 1	Traditional clothing in Saudi Arabia	Descriptive	130 words/10 lines	Present tense
Text 2	True friends	Descriptive/narrative	116 words/8 lines	Present/past tense
Text 3	Sport	Descriptive/narrative	109 words/8 lines	Present tense
Text 4	Arabian coffee	Descriptive	131 words/9 lines	Present tense

After the study participants completed the four reading sessions, they took an immediate post-test that was identical to the pre-test. However, sentences appeared in a different random order. Details of the study tests are discussed in the following section.

4.4.2.2 Testing materials

The pre-test, post-test, and delayed post-test consisted of the same multiple-choice recognition task that contains two parts and applies to the same words and uses the same test items (see Appendix E). The first part tested the participants on the *word form* by providing alternative versions of the word with different vowels but the same consonantal structure. The alternative versions are either existing English words or non-

words. An alternative existing word is chosen when it has an identical consonantal structure to the target word (e.g., sweet and sweat) otherwise a non-word is made by keeping the same consonantal structure and only changing the vowels (e.g., steps and staps). The second part tested the participants on the *word meaning* by displaying three different Arabic translations of the English word (see Figure 4.4). The target words are embedded in a sentence to provide context and meaning cues in order to ensure that the target word was uniquely identified. This decision was prompted by previous research. Unlike previous studies on vowel blindness which mostly examined the decoding of words in isolation rather than in context (Alsulaimani, 1990; Hayes-Harb, 2006; Ryan & Meara, 1991; Saigh & Schmitt, 2012; Stein, 2010; Taylor, 2008), Hayes-Harb (2006), for instance, suggests that data collected from contextual reading could provide more realistic information on the vowel blindness phenomenon as it more closely resembles the regular mode of reading.

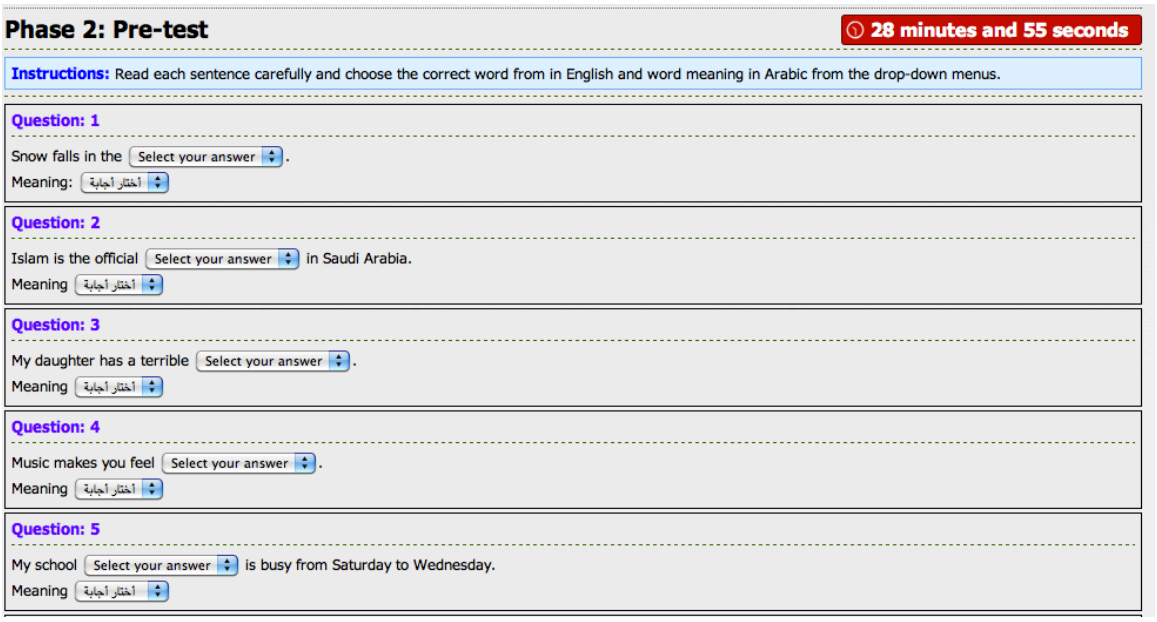


Figure 4.4. An example of the multiple-choice recognition test

The 40 test items, i.e., multiple choice questions of each of the three tests contained 32 target words and 8 nontarget distractors. While the format was the same for all three tests, the order of the test items was separately randomized for each test.

4.4.2.3 Test words

The study concerns two types of words: treated words (targeted) and untreated words (nontargeted). The treated/targeted words are 32 words that were embedded in the four reading passages of the study and for which help options were provided, except in the control condition. The untreated/nontargeted words are the distractor items that were also tested and were used to test the transferability effect of the treatment to untreated words. However, these untreated words appear only in the tests and never in the reading passages. The test words were selected by considering word frequency, word length, and syllable type.

High-frequency words have mainly been examined in vowel blindness studies (Hayes-Harb, 2006; Ryan & Meara, 1991; Saigh & Schmitt, 2012; Taylor, 2008) because researchers agree that vowel blindness is primarily a word form problem with words that are familiar in meaning to the learner (see section 2.1 and section 2.2). Saigh and Schmitt (2012) argue that if the problem exists at the level of high-frequency words, then it will definitely be harder to examine it with mid- or low-frequency words without avoiding the influence of other variables such as learnability of the words and frequency of exposure.

Vowel blindness data from high-frequency words provides an interesting and revealing selection of vowel blindness errors in decoding and encoding English words. Accordingly, the researcher created a database of 264 vowel-blindness errors in decoding English words from previous work of Alsulaimani (1990); Bowen (2011); Ryan and Meara (1991); Saigh and Schmitt (2012). These errors were made by ESL/EFL readers from various backgrounds, including Saudi, Omani, and Emirati readers, as well as readers from unidentified Middle-Eastern backgrounds. From this pool of 264 words, the study test words were selected by taking into account word length and syllable type.

As for word length, Ryan and Meara (1991), for instance, emphasize its importance in selecting words for vowel-blindness studies because they found that even non–vowel-blinded ESL learners are overwhelmed in processing longer words. Most commonly, studies (Hayes-Harb, 2006; Ryan & Meara, 1991) limit the length to 10

letters per word. For this reason, the current study included test items with a maximum of 10 letters per word to ensure compatibility with previous research.

Finally, syllable type is another important factor in test word selection, especially because the current study employs syllable-focused glosses. Syllabification is a method of teaching phonics to L1 English children. This approach emphasizes learning reading skills by decomposing words into syllables based on the vowel information contained in each syllable. The vowel information determines the syllable type as open, closed, vowel team, consonant-vowel-e, vowel-r, schwa syllable, and consonant-le (Gates & Yale, 2011). Mixing different syllable types is generally encouraged in vocabulary learning because it eliminates the same-pattern learning effect (Laufer & Hill, 2000). For this reason, the current study included words with a variety of syllable types and with words up to four syllables so as to examine fairly the effect of the syllable-focused help option.

4.4.2.4 Help treatment conditions

As stated before, VALE provided four form-focused help treatment conditions and one control condition. All these conditions provided the same reading passages, word meaning reading exercises, and word meaning options with uniform page features and design including the timer, the page format, and font size and colour. The only difference between the conditions came from the presence or not of form-focused help of some sort (see Figure 4.3 on p. 77 and Figure 4.5 on p. 82). The two main help options employed in the study are input enhancement, which was added to the displayed reading text, and form-focused glosses, which, on a word in the text being clicked, appeared in the bottom left hand corner of the page. Within the form-focused glosses, there are three different conditions: segment-focused glosses, syllable-focused glosses, and segment-syllable focused glosses. The focus of all of these help options is to provide form-related vowel information in order to reduce vowel blindness for Arabic ESL learners. The word-meaning help information was common to all the treatments in the study, including the control group (see Appendix G for a screen shot of the control group), namely the optional access to word meaning help in the participants' L1, which was available to enable them to successfully complete the accompanying exercise on word meaning. As illustrated in Figure 4.5, the input enhancement condition has the vowels highlighted in yellow with the whole word underlined to indicate a clickable link,

which in this treatment condition is only to the word-meaning information. Also, the hand cursor appears once the learner points at the word to enable the hyperlink-clicking function.

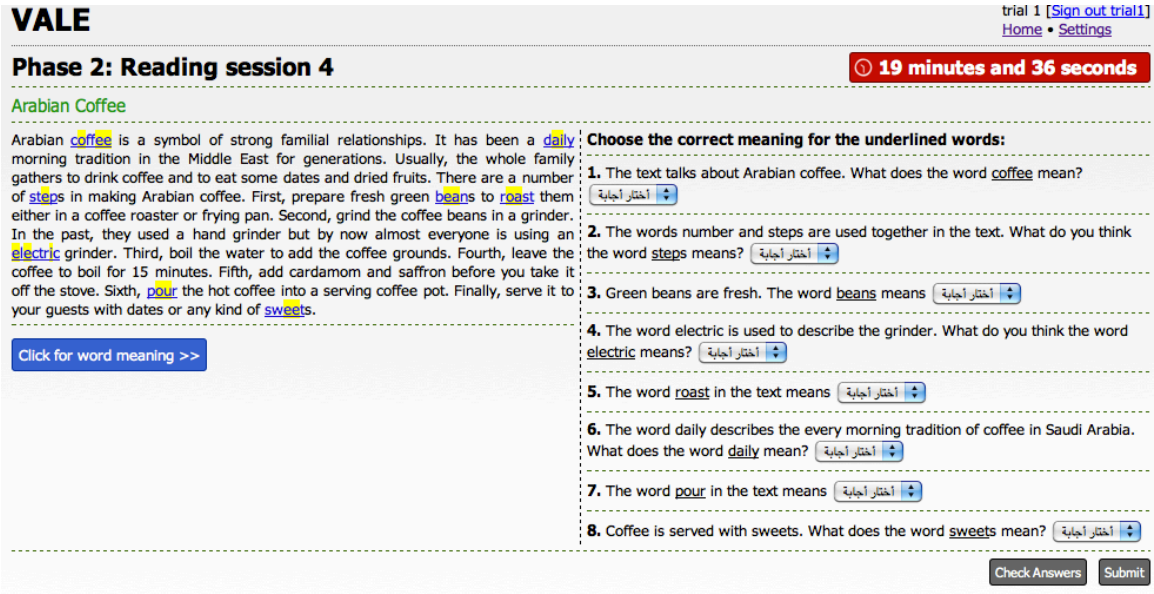


Figure 4.5. Input enhancement condition showing the link to the optional word-meaning information

In the conditions with form-focused glosses, the link to the word-meaning option appears via the same hyperlink as for the input enhancement condition, but this is presented at the bottom of the form-focused gloss, not at the end of the reading text. The following section explains each experimental condition and provides more details about the content of the form-focused glosses.

Input enhancement

One way to bring the attention of Arabic EFL learners to vowels is through input enhancement. Figure 4.5 illustrates how input enhancement is achieved in VALE through highlighting. In the reading passage about *Arabian coffee*, VALE underlines the target word and highlights the vowel(s) in yellow. The decision to underline the whole word is to make it stand out for the reader as a clickable link for optional access to the word meaning. This clickable format was controlled across all the study conditions including the control group by using the same hyperlinking format for all the groups.

According to Sharwood Smith (1993), the main function of the input enhancement technique is to draw the learner's attention to linguistic form by enhancing its saliency through typographical cues such as bolding, italicizing, capitalizing, underlining, changing the size, font, and colour; or by highlighting as discussed in section 3.3.2. Based on previous research (see chapter 2, p. 8), English has a deep orthography which is found to be visually accessed and processed rather than phonetically by Arabic ESL learners (Bowen, 2011). Vowel blindness is found to occur due to interference of processing only consonants while vowels are contextually inferred (Fender, 2008; Randall & Meara, 1988; Ryan & Meara, 1991). Research by Alsadoon and Heift (2015), for instance, found that input enhancement, namely in the form of highlighting, is effective in attracting Arabic ESL learners' attention to English vowels during short sentence reading tasks.

Simard (2009) shows that the format of the typographical cues affects noticing and intake differently. Simard (2009) and Farahani and Sarkhosh (2012) report that there are input enhancement formats that have inherent saliency. One example is highlighting, which increases the chances of the learners noticing the target form. Therefore, the choice of highlighting as the typographical cue to use in the study is based on this previous research (Farahani & Sarkhosh, 2012; Simard, 2009) (see section 3.3.2) and uses contrast of colours, specifically yellow highlighting of the vowels against the black color of the surrounding text, and the white background of the screen. Highlighting is used to make the learner notice the written form of the vowel. However, it contributes only to learning the written form of the word since it does not supply any information about vowel sound either as audio, or transcription, or grapheme-phoneme correspondences like the other form-focused help options.

Segment-focused glosses

The segment-focused glosses in the current study provide, amongst other things, information on the rules of grapheme-phoneme correspondences. More specifically, VALE provides the following information as illustrated in Figure 4.6 with the word *wool*: the pronunciation of the word, the part of speech; the IPA transcription; the vowel graphemes; the IPA symbols of the vowels; and the pronunciation of the vowels.

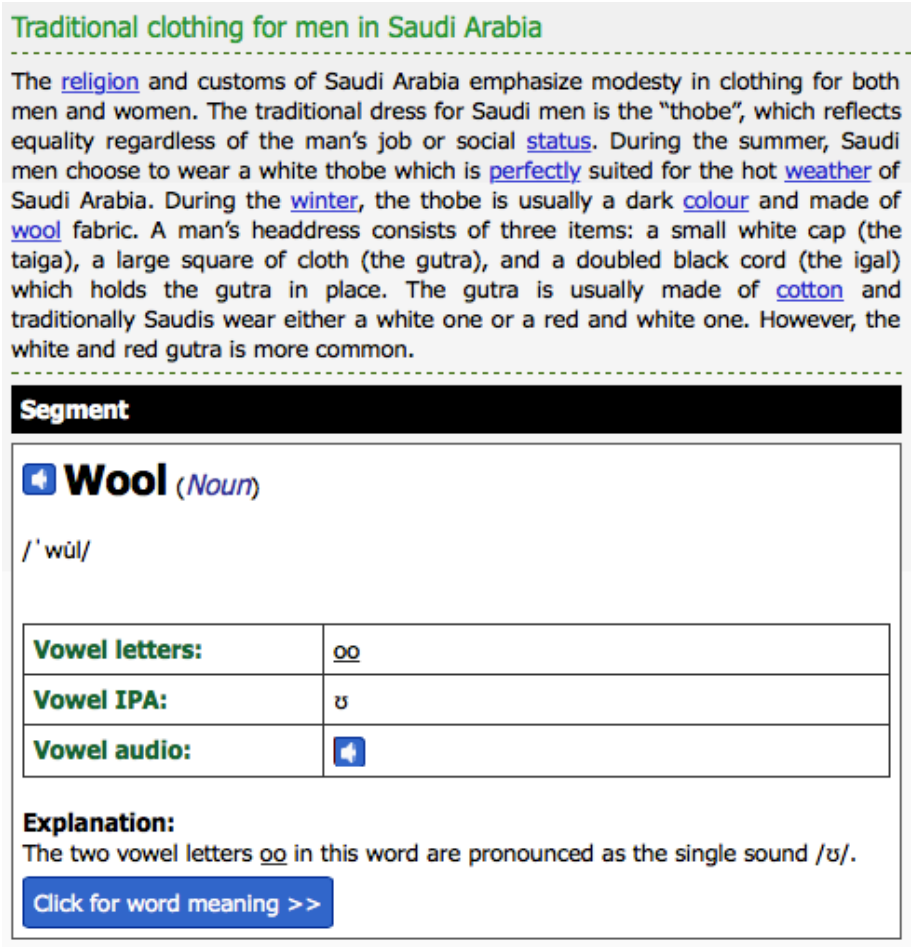


Figure 4.6. An example of the segment-focused glosses

Based on the orthographic-depth hypothesis (see chapter 2, p. 14), English orthography is considered to be an opaque language with respect to grapheme-phoneme mapping as it does not approximate a one-to-one letter-to-sound correspondence. ESL/EFL learners from languages with more transparent writing systems such as Arabic need to have phonological (i.e., phoneme-grapheme) awareness in order to decode and encode an opaque L2 orthography (Nadia & Charles, 2011). Fender (2003) and Taylor (2008) recommend training in phonological awareness for Arabic EFL learners to develop better vowel decoding skills. Thus, the study provides help in the form of phonics-type segment information in glosses that focus on the vowels' grapheme-phoneme mappings.

Moreover, Taylor (2008) stresses the importance of helping Arabic EFL readers become more aware of the grapheme-phoneme correspondence patterns, namely through segment phonics teaching. Generally, Groff (1971) notes that *vowel patterns* is the common term that is used to describe the general rules that govern the grapheme-phoneme mappings of vowels. There are five vowel patterns in English: 1) short vowels; 2) long vowels; 3) schwa; 4) diphthongs, and 5) vowel digraphs (i.e., two letters representing one vowel sound) (Gates & Yale, 2011).

Along with showing the correspondence of vowel letters with sounds, segment-focused glosses generally provide pure phonological information in the form of vowel and word audio recordings and IPA transcriptions. Research by Fender (2003), Naser (1963), Saigh and Schmitt (2012), and Taylor (2008) recommends explicit training of Arabic EFL learners in both pure and orthographic phonological awareness that includes vowel pronunciation, the graphemic form of the vowels, as well as phonics teaching (i.e., the link between vowel grapheme and phoneme).

Syllable-focused glosses

With syllable-focused glosses, the current study adopts the syllabification approach for one of the form-focused glosses conditions. This approach is used in teaching reading skills to native English children (Fox & Routh, 1975; Groff, 1971, 1981) and L2 learners (Stein, 2010). It teaches reading by identifying syllables in written words as the major cue for word recognition. Groff (1981), for instance, promotes the teaching of syllable-phoneme sequences for beginning readers before letter-phoneme correspondences.

Syllable-focused glosses decompose words into syllables and focus on the vowel information to predict the type of syllable. Specifically, the glosses in VALE include division of the word into syllables, the type of syllables, and finally, general explanations about syllable rules. Figure 4.7 gives an example of syllable-focused glosses illustrating the word *status* which is divided into two syllables. The first syllable is open ending with *a* while the second syllable is closed consisting of a vowel *u* followed by consonant *s.* The rules for syllable types are represented under the table followed by the usual optional access to word meaning in Arabic.

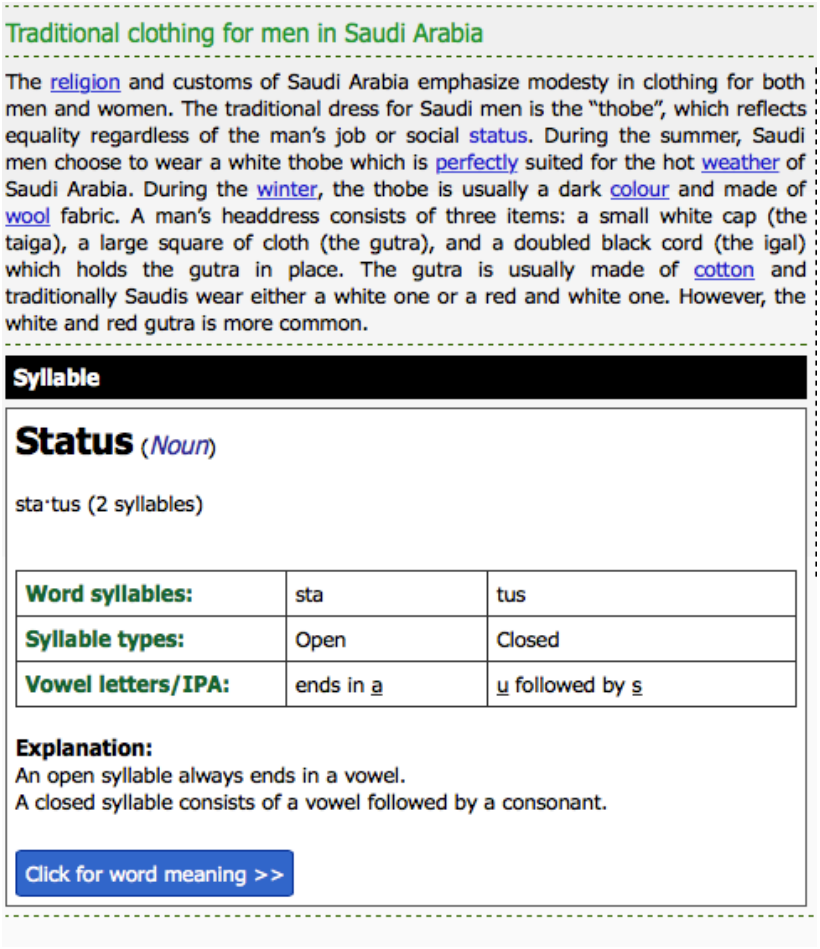


Figure 4.7. An example of the syllable-focused glosses

The syllabification approach has been chosen for three reasons:

Firstly, and, based on the hypotheses and findings discussed in section 2.4.2.1, learners from Semitic language backgrounds have a tendency to identify words through decomposition of words into roots and affixes (Abu-Rabia & Awwad, 2004; Lukatela et al., 1980; Saiegh-Haddad & Geva, 2008; Taft, 1981). Therefore, Arabic ESL learners are found to transfer their L1 habits of decomposing the word into smaller parts when reading English (Hayes-Harb, 2006; Ryan & Meara, 1991). Even though morphemes do not equate one-to-one with syllables either in Arabic or English, the syllabification approach is thought to facilitate reading English words by providing something closer to Arab learners' cognitive habits of decoding. Bowen (2011) encourages the syllabification

approach as it helps in breaking the word into smaller and easier parts for decoding and encoding.

Secondly, vowels in English syllables can be affected by the consonantal patterns preceding or following them. Treiman and Kessler (2005) confirm that these consonantal patterns help native readers of English to disambiguate the target vowel. “For example, English /ɛ/ is more likely to be spelled as ea before d (e.g., head, instead) than before other consonants (e.g., set, west). By no means is every /ɛ/ before d spelled with ea—bed and wed do not use this spelling—but a child who knew the association would be a better speller than a child who did not” (Treiman & Kessler, 2005, p. 25). Accordingly, these examples reveal that grapheme-phoneme correspondences are sometimes easier to state for syllables than segments. Based on vowel blindness research by Stein (2010), Arabic ESL readers lack sensitivity in noticing vowel-consonant associations at the syllable level.

Thirdly, there are various methods for teaching reading and spelling. These are usually referred to as phonics. Example of phonics approaches to reading and spelling are phonograms, word families, spelling patterns, graphemes, and syllabification (Groff, 1981). Most of these methods are similar except in the number of vowel-consonant association patterns. The similarity of all of these methods is the use of vowel cues to identify word consonant associations, syllables, and the word’s entire structure for word identification in reading.

The current study implements the syllabification method, in which the vowel information determines the kind of syllables in the word. The choice of this method specifically was made on logistical grounds. This method is simpler than the other phonics methods with respect to the number of patterns and the number of vowel-consonant associations. For instance, the phonograms method generally recognises seventy vowel-consonant associations while syllabication follows a rule of six common syllable types. Given the number of target words and the time constraints of the study, it was more plausible to use a simple and minimalist method to make it as easy as possible for the learners to process the additional information of the glosses.

Segment-syllable focused glosses

The third kind of form-focused condition is segment-syllable glosses, which include a combination of information from both segment and syllable glosses. The information that was selected from both glosses is the following: 1) the division of the word into syllables; 2) the types of syllables; 3) the pronunciation of the word, 4) the IPA transcription of the word, 5) the vowels as letters; 6) the IPA symbols of the vowels; and 7) the explanation about the syllable types. Figure 4.8 provides an example of a segment-syllable focused gloss.

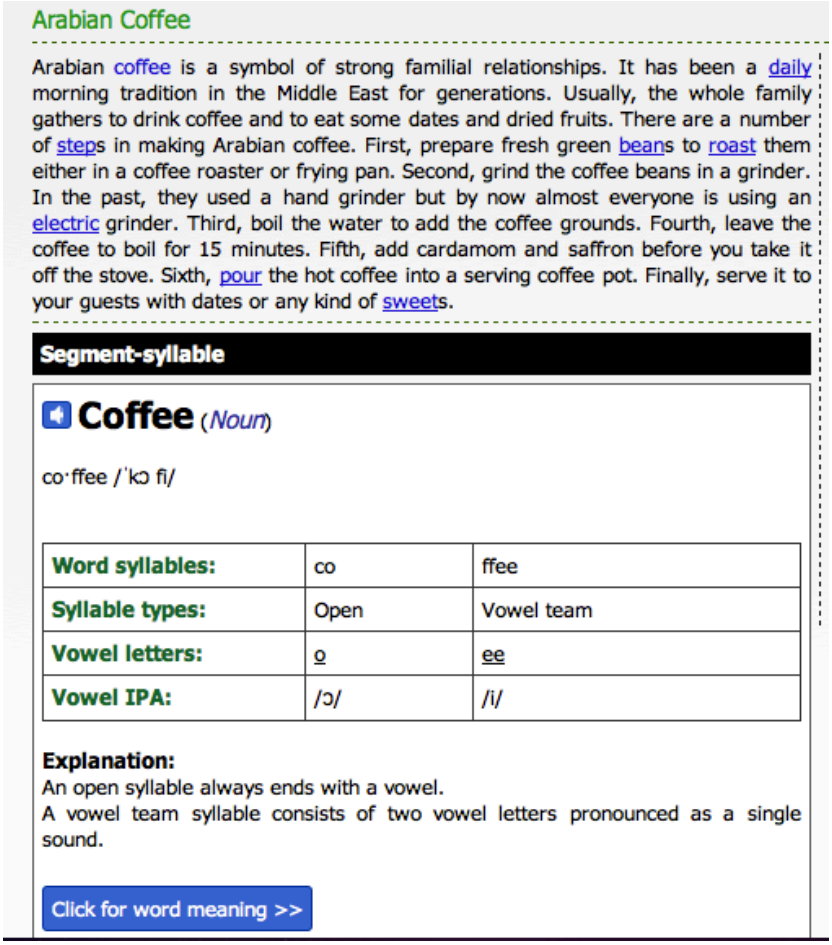


Figure 4.8. An example of the segment-syllable focused glosses

The decision to use a combination of the syllable and segment information is simply to test if the learners' vowel blindness can be removed best by combining together the benefits of both separate types of help. According to Taylor (2008), Bowen

(2011), and Abu-Rabia and Awwad (2004), both awareness of grapheme-phoneme mappings and awareness of syllables are essential for word recognition especially for learners with a Semitic language background. According to Johnston, McGeown, and Watson (2012) and Groff (1981), the blended phonics approach is a popular trend in teaching reading for English native children: this approach moves from teaching segment correspondences to segment sequences and syllable correspondences up to the whole word.

Main justifications for the study's help options

Table 4.4 presents a summary of the main arguments used to justify the choice of the four form-focused help options in VALE. They are based on the nature of the English writing system, the Arabic writing system, L1 word recognition processes, and L2 teaching methods for English word reading.

Table 4.4. Summary of the main arguments for the help options in VALE

Help options	English (L2) writing system	Arabic (L1) writing system	L1 word recognition process	Teaching method for English word reading
Input enhancement	It represents consonants more consistently than vowels	It is generally classified as a consonantal writing system with vowels partly unrepresented	Consonantal information is visually processed while vowel information is inferred from context (Fender, 2008; Ryan & Meara, 1991) L1 visual processing of word form is transferred (Bowen, 2011; Randall & Meara, 1988)	Input enhancement is generally used for attracting visual attention to word form in L1 and L2 (Sharwood Smith, 1993) Also, it is found to effectively improve orthographic vowel knowledge for Arabic ESL learners (Alsadoon & Heift, 2015)

Help options	English (L2) writing system	Arabic (L1) writing system	L1 word recognition process	Teaching method for English word reading
Segment	A pure alphabet would work on a one grapheme to one phoneme basis but English writing departs quite a lot from that so segmental grapheme-phoneme correspondences are complex (Nadia & Charles, 2011)	It is often described as a transparent orthography with one-to-one grapheme-phoneme correspondence (Nadia & Charles, 2011)	L1 studies explaining L2 vowel blindness suggest that in the L1 Arabs rely on reading consonantal segments rather than vowel segments (Abu-Rabia, 1995; Abu-Rabia & Taha, 2004)	Phonics teaching in the form of grapheme-phoneme correspondences is widely used in teaching of English native children and strongly recommended for Arabic ESL/EFL learners (Taylor, 2008)
Syllable	Grapheme-phoneme correspondences are sometimes easier to state for syllables than segments (Treiman & Kessler, 2005)	It sometimes can appear like a syllabic system where one letter stands for a syllable, as in Figure 2.3 *faʕala*. Also prefixes and suffixes often closely correspond to syllables, as in *alkalbu* (i.e., the dog)	Other research suggests that Arabs decode through decomposition of word into smaller parts (i.e., roots, affixes, and syllables) in L1(Abu-Rabia & Awwad, 2004; Saiegh-Haddad & Geva, 2008)	Phonics teaching in the form of syllabification is commonly used for teaching of English native children and recommended for Arabic ESL/EFL learners (Bowen, 2011)
Segment-syllable	Both grapheme-phoneme and grapheme-syllable are essential for orthographic word recognition (Abu-Rabia & Awwad, 2004; Taylor, 2008)	Transparent orthography with some syllabic features	Grapheme-phoneme and syllable processing is essential for orthographic recognition (Bowen, 2011; Taylor, 2008)	Blended phonics is the most commonly taught version of phonics to English native children (Groff, 1981)

4.4.3 Post-treatment phase

The post-treatment took place 6 weeks after the completion of the main treatment phase. Study participants completed a delayed post-test (identical to the pre-

test and post-test but with a different randomised order of items) and filled in an attitude questionnaire through VALE. This phase lasted 60 minutes. Finally, 10 students from each experimental group (i.e., 40 in total) were randomly selected to complete semi-structured retrospective interviews which lasted an average of 15 minutes.

After finishing the delayed post-test, an attitude questionnaire evaluated the participants' perception of their learning experience in VALE in general and the ease and usefulness of its help options (see Appendix D). The attitude questionnaire was written in a Likert scale format and consisted of 26 questions (see Figure 4.9). All participants received the same general questions about VALE. However, questions relating to the help options were geared to the help options each learner had received, or no help options as was the case for the control group. The questionnaire generally took 30 minutes to answer.

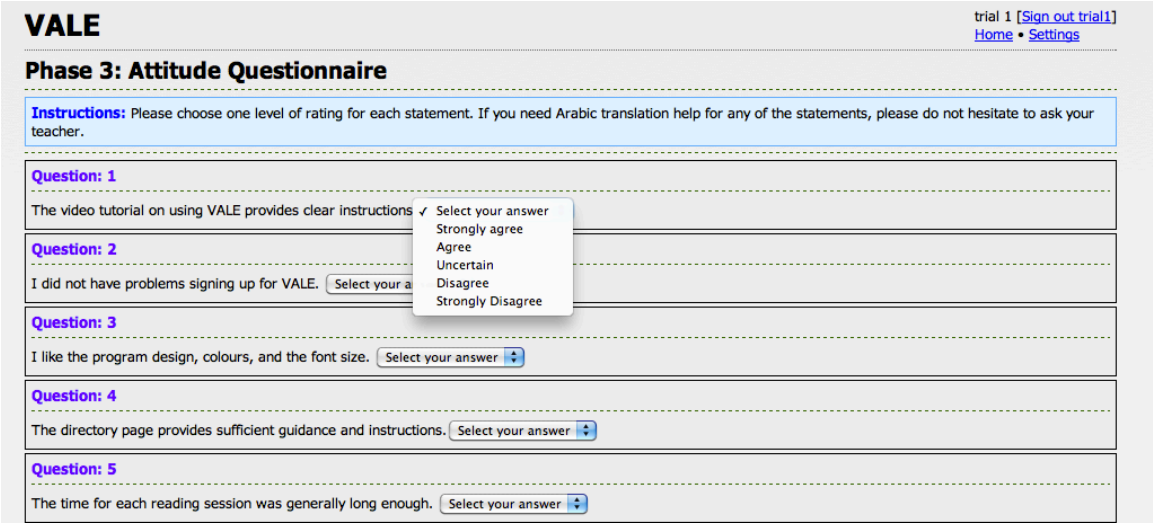


Figure 4.9. Sample of the attitude questionnaire

Once the study participants submitted the attitude questionnaire, a final thank you message was displayed to indicate the end of the study (see Figure 4.10). For 10 participants, the thank-you message came with a prize of 50 R.S. which the program randomly assigned to 10 participants.

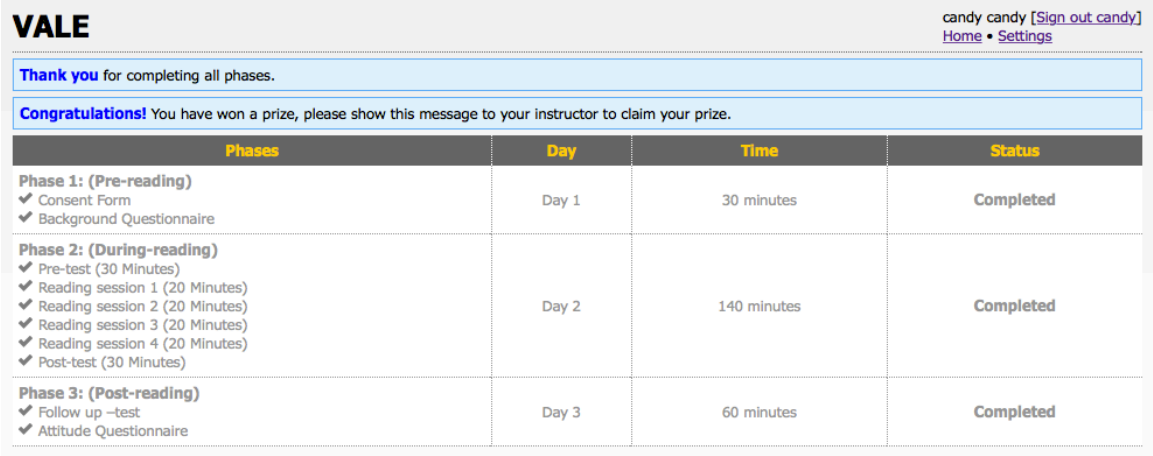


Figure 4.10. Thank-you and winning message

4.4.3.1 Retrospective interviews

Retrospective interviews are commonly used in noticing studies in SLA. For instance, Izumi and Bigelow (2000) performed retrospective interviews on four out of the nine participants in their study to examine the role of L2 production in promoting noticing. The authors confirmed that the interviews were beneficial in that learners reflected on the perception of the problems in their production and the interviews' relevance to the noticing of their input. Leow (2007) also stresses the importance of using qualitative methods to ensure internal validity in studies of noticing and cognitive concepts.

In the current study, ten participants from each group, that is, the four experimental groups, were asked 15 questions (see Appendix F) which, however, in some cases required follow-up questions for clarification. The interview data were meant to augment attitude questionnaire data and provide qualitative accounts of participants' awareness of vowel blindness. The participants were given a choice of responding either in their L1 or the L2, to ensure that they were able to express their thoughts clearly. In the event, all chose to speak in Arabic. An Olympus digital recorder was used and the data were transcribed right after the completion of the interviews.

The retrospective interviews were designed to be both open-ended and semi-structured to allow the researcher to ask for clarifications and elicit additional

information[4]. They further provided information on whether the learners' awareness and attitudes towards the impact of VALE was consistent with the quantitative results of the study.

4.5 Data analysis

The study gathered two types of data: quantitative and qualitative. The quantitative data relates to the first two sets of research questions and in part to the third (the attitude questionnaire) while the qualitative data answers the third set of research questions (see section 4.1).

4.5.1 Quantitative data

The first two sets of research questions required measurement of the effectiveness of VALE's help options based on changes in vowel blindness scores between the pre-test, post-test and delayed post-test. When scoring students' responses, a correct answer was given 1 point whereas an incorrect answer was given 0 for each part of the test question, that is, separately for form and meaning. Vowel blindness was evidenced when the learner received 0 points for the word form and 1 point for the word meaning.

As for the inferential statistics applied to the quantitative data in the form of test scores, Table 4.5 displays the research questions of each set along with the statistical tests that were applied to the data.

[4] Initially, the study planned to rely on think-aloud protocols in an attempt to access the participant's thoughts during the main treatment. However, the participants in the pilot study in 2014 rarely expressed themselves aloud. Study participants reported that thinking aloud distracted them from focusing on the study's main tasks. They also reported that they had little idea of what to say and that they needed some cues to guide their thoughts. For this reason, retrospective interviews were adopted to collect data more successfully.

Table 4.5. Research questions and statistical procedures

Research questions set 1:	Statistical effect tested	Inference test
RQ 1.1 Over the initial intervention period (pre-test to post-test) do four different help options (in the form of input enhancement and three types of form-focused glosses) reduce vowel blindness more than no help at all? Do the four options differ in their effects on vowel blindness reduction?	Change (decrease/increase) from pre-test to post-test for treated words	2x5 mixed ANOVA
	The differences between the experimental groups in amount of change	ANOVA with a Tukey Post Hoc
RQ 1.2 Over the retention period (post-test to delayed post-test) do the four different help options lead to maintenance of any reduction in vowel blindness better than no help at all? Do the four options differ in their effects on maintenance of vowel blindness reduction?	Change (decrease/increase) from post-test to delayed post-test for treated words	2X5 mixed ANOVA
	The differences between the experimental groups in amount of change	ANOVA with a Tukey Post Hoc
RQ 1.3 Over the entire period of the study (pre-test to delayed post-test) do the four different help options yield a net reduction in vowel blindness more than no help at all? Do the four options differ in their effects on overall vowel blindness reduction?	Change (decrease/increase) from pre-test to delayed post-test for treated words	2x5 mixed ANOVA
	The differences between the experimental groups in amount of change	ANOVA with a Tukey Post Hoc
Research questions set 2:	**Statistical effect tested**	**Inference test**
RQ 2.1 Over the initial intervention period (pre-test to post-test) do four different help options (in the form of input enhancement and three types of form-focused glosses) reduce vowel blindness of nontarget/untreated items more than no help at all? Do the four options differ in their effects on vowel blindness reduction of nontarget/untreated items?	Change (decrease/increase) from pre-test to post-test for untreated words	2x5 mixed ANOVA
	The differences between the experimental groups in amount of change	ANOVA with Tukey Post Hoc

Research questions set 2:	Statistical effect tested	Inference test
RQ 2.2 Over the retention period (post-test to delayed post-test) do the four different help options lead to maintenance of any reduction in vowel blindness of nontarget/untreated items better than no help at all? Do the four options differ in their effects on maintenance of vowel blindness reduction of nontarget/untreated items?	Change (decrease/increase) from post-test to delayed post-test for untreated words	2x5 mixed ANOVA
	The differences between the experimental groups in amount of change	ANOVA with Tukey Post Hoc
RQ 2.3 Over the entire period of the study (pre-test to delayed post-test) do the four different help options yield a net reduction in vowel blindness of nontarget/untreated items more than no help at all? Do the four options differ in their effects on overall vowel blindness reduction of nontarget/untreated items?	Change (decrease/increase) from pre-test to delayed post-test for untreated words	2x5 mixed ANOVA
	The differences between the experimental groups in amount of change	ANOVA with Tukey Post Hoc

Table 4.5 indicates that the statistical procedures for the two sets of research questions require repeated use of 2x5 mixed ANOVAs to test the effectiveness of the help options in reducing vowel blindness comparing all the study groups from pre-test, immediate post-test, to delayed post-test. Tukey post hoc tests were used to measure the differences among the experimental groups in vowel blindness reduction.

Quantitative data arising from the attitude questionnaire relevant to the third research question were simply summarised for each group with descriptive statistics.

4.5.2 Qualitative data

The qualitative data addressed the third set of research questions which concern the learners' awareness of the vowel blindness problem before and after the study as well as their attitudes towards their learning experience with VALE. The data for RQ 3.1 and 3.3 were collected from the attitude questionnaire and the retrospective interviews while RQ 3.2 was answered solely from the retrospective interviews.

The data of the retrospective interviews were coded for the effect on vowel blindness of help options with respect to three major themes: 1) the learners' evaluation

of the technical design of VALE, 2) the learners' awareness of the vowel blindness problem, and 3) the learners' perception of the learning experience through VALE's help options.

After the qualitative analysis, the coded data were further analyzed quantitatively. Descriptive statistics were calculated as frequency counts and percentages to show differences in positive and negative opinions between different help option groups.

4.6 Summary

This chapter outlined the research methodology by presenting the three sets of research questions that underlie the current study. The first set of research questions investigates the effect on vowel blindness of VALE's help options for targeted words. The second set examines the effect on vowel blindness of VALE's help options for untreated words. The third set investigates the study participants' awareness of vowel blindness and attitudes towards VALE. The chapter further described its study participants, the mixed-method between groups research design with its procedure, intervention, and instruments largely administered through VALE. Finally, it detailed the data collection and analysis of both quantitative and qualitative data. The results of the study are presented in the following chapter.

5 Results

This chapter presents the results which address the study research questions of this dissertation. The chapter is therefore divided into three main sections corresponding to the three sets of research questions: the effect on vowel blindness of VALE's help options for targeted words, the effect on vowel blindness of VALE's help options for untreated words, and the study participants' awareness of vowel blindness and attitudes towards VALE. The findings are derived from two types of analyses: 1) quantitative (i.e., pre-test, post-test, and delayed post-test scores), and 2) qualitative (i.e., interviews and questionnaires).

5.1 Research topic 1: Vowel blindness and help options

This research topic concerns the effect of four different help options on reduction of vowel blindness by asking three research questions: 1) RQ 1.1, in section 5.1.2, tests the effect of the different form focused help options on reducing vowel blindness in targeted words between pre-test and post-test, 2) RQ 1.2, in section 5.1.3, examines the effect of the different help option treatments on retention of any improvement in vowel blindness on targeted words after 6 weeks from post-test to delayed-post test, during which time there was no further help, and 3) RQ 1.3, in section 5.1.4, investigates the effect of the different help option treatments on reducing vowel errors in targeted words over the entire study period from pre-test to delayed post-test.

The results for these questions are provided through the analysis of the changes in the participants' scores from the pre-test to post-test and then from post-test to delayed post-test. The 32 test items explicitly targeted by the help treatments are therefore included in this analysis, out of 40 test items in all. Distractors (8 words) are analyzed for the next research topic to test the participants' ability to transfer the help option knowledge to new words (see section 5.2). The word form error scores on items

whose meaning is known as determined by the participant's score on word meaning (referred to below for short as 'word form error') are used as the primary measure of vowel blindness (see section 4.5.1): the analyses examine the effects of the treatments, which in the current study were implemented in independent groups of participants, on the changes (decrease/increase) in these scores. The study hypothesis is that there will be a decrease in word form errors due to the effect of help option treatments. The null hypothesis is that there will be no difference between the effects of the different treatments, including no help at all (control) (see section 4.5.1).

Before answering this research topic's questions, an analysis of the baseline pre-test scores is performed to show that all the participants are starting at a similar level by measuring whether all the study groups have a similar mean score and distribution for word form prior to receiving the study's different help option treatments.

5.1.1 Pre-test analysis

Table 5.1 provides descriptive statistics for the control group, the four experimental groups (i.e., syllable, segment, segment-syllable, highlighting), and the means for word form errors.

Table 5.1. Descriptive statistics for word form errors in treated words in the pre-test

Study groups	Word form				N
	Mean	SD	Minimum	Maximum	
Control	21.31	2.54	19	27	50
Syllable	21.64	2.63	19	27	50
Segment	20.82	2.50	19	27	50
Segment-syllable	21.24	2.22	19	26	50
Highlighting	21.14	1.98	19	26	50
Total	21.23	2.38	19	27	250

Table 5.1 shows that the participants selected for the study scored an overall mean of 21.23 for word form errors out of 32 test items (i.e., 66.3%), with a standard deviation of 2.38. The means are quite similar for each group. The descriptive statistics

therefore imply that the groups start off at a reasonably equal level and hence that the experiment has successful control over the participants' baseline level with respect to vowel blindness.

Inferential statistics were then used to test the null hypothesis that all groups have the same mean score. A one-way between groups ANOVA with an alpha level of 0.05 is used even though the data are found to violate the assumption of normal distribution due to a more spread out distribution (see Appendix H for each group's skewness and kurtosis). However, ANOVA is generally argued to be robust against violations of the normality prerequisite. This was confirmed by the Kruskal-Wallis test which gave very similar results to ANOVA (see Appendix H).

Table 5.2 reports the ANOVA results for pre-test word form error scores. The findings show that the groups do not significantly differ in word form errors: $F(4, 245) = .77$, $p = .54$. Therefore the null hypothesis is confirmed at the 95% level of significance and all groups started at a very similar level of knowledge with respect to our crucial measure of vowel blindness.

Table 5.2. ANOVA results for word form errors in the pre-test in treated words

Scoring categories	Source	Sum of Squares	Df	Mean Square	F	P
Word form	Between groups	17.577	4	4.394	0.768	0.547
	Within groups	1407.02	245	5.720		
	Total	1424.59	250			

In sum, the pre-test confirms that, following our participant selection procedure, all the groups started at an equivalent level for vowel blindness (i.e., wrong selection of word form for words whose meaning is known).

5.1.2 RQ 1.1: Over the initial intervention period (pre-test to post-test), do four different help options (in the form of input enhancement and three types of form-focused glosses) reduce vowel blindness in targeted items more than no help at all? Do the four options differ in their effects on vowel blindness reduction?

To answer RQ1.1, the results from pre- and post-test need to be compared to examine the score changes over time from pre-test to post-test, and the effect of the interaction between the study groups and time. That is, do the experimental groups yield greater reductions in vowel blindness than the control group? And which experimental group has the strongest vowel blindness reduction effect?

Accordingly, a 2x5 mixed ANOVA was performed with time (pre-test vs. post-test) and study group (4 experimental and 1 control) as the independent variables and each test measure in turn as the dependent variable.

Table 5.3. 2X5 Mixed ANOVA of word form error changes from pre-test to post-test in treated words

Tests of within-subjects effects						
Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Time	13077.49	1	13077.49	2850.43	.000	.921
Time*group	3561.19	4	890.29	194.054	.000	.759
Error (Time)	1128.62	245	4.588			
Tests of between-subjects effects						
Intercept	130545.36	1	130545.36	11542.08	.000	.979
Group	3784.530	4	946.133	83.652	.000	.576
Error	2782.354	245	11.310			

The results for the word form error scores displayed in Table 5.3 reveal a significant difference across the two time points (when all participants are considered together), $F(1, 245) = 2850.43$, $p<.001$ with an effect size of .921 (partial eta squared). This means that word form scores have changed significantly over time, suggesting a real overall change in the vowel blindness level from pre-test to post-test. The between subjects effect in Table 5.3 also shows a main effect for groups, $F(4, 245) = 83.652$, $p=<.001$ with an effect size of .576 (partial eta squared), implying significant variance in

the learners' performance between the study groups when pre-test and post-test scores are considered together. Most importantly, however, a significant interaction is also found between time and groups, $F(4, 245) = 194.054$, $p<.001$ with an effect size of .759 (partial eta squared), meaning that the change in the word form scores from pre-test to post-test was different in different study groups.

In order to reveal whether the experimental groups differed from the control group or from each other in their effect on vowel blindness, Tukey multiple comparisons between the study groups of pre-post change scores were next carried out (see Appendix I for Tukey's multiple comparisons). Figure 5.1 displays two graphs which provide the means of the vowel blindness errors for all groups for the pre-test and post-test, respectively.

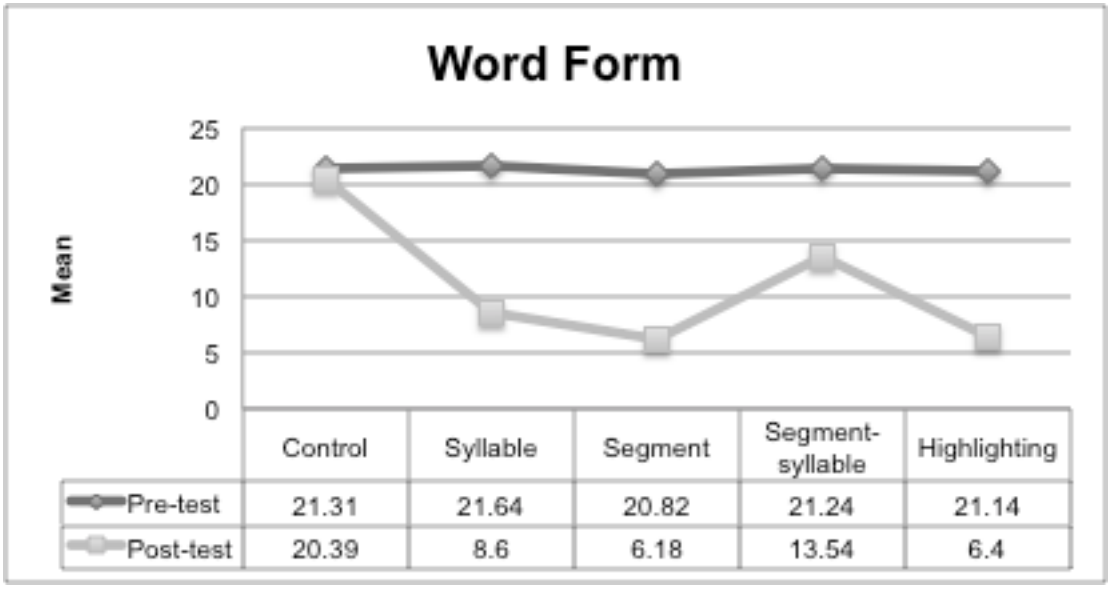


Figure 5.1. Change in vowel blindness errors from pre-test to post-test in treated words

Table 5.4 shows the mean reduction rate of vowel blindness errors between pre-test to post-test. The results reveal three notable findings:

Firstly, significant differences were found between the control and all the other groups with a p value of <.001 (see Appendix I). Figure 5.1 and Table 5.4 show that there was almost no reduction (<1 item on average, 3%) in vowel blindness errors for the

control group while all the experimental groups showed some reduction, of at least 7 out of 32 items (24%).

Secondly, a significant difference was found between the segment-syllable help treatment group and all the other groups with a p value of <.001 as presented in Appendix I. Figure 5.1 and Table 5.4 indicate that this is due to this treatment not yielding as great a vowel blindness reduction of 7 items on average (24%) as the other three experimental treatments, though still significantly better than no help treatment at all. Possible reasons for this are discussed in section 6.2.3 of the Discussion chapter.

Thirdly, the highlighting and segment treatments do not differ significantly from each other (p<.001) in the Tukey paired comparisons presented in Appendix I and Table 5.4. Moreover, the difference in segment and syllable treatments is also nonsignificant at p=.065. However, a significant difference is found between highlighting and syllable treatments with p=.043 as given in Appendix I. This shows that, at least in the initial intervention phase of the study, these types of help options generate a remarkable improvement of at least 13 out of 32 items (41%) whose form is no longer wrongly chosen, almost twice the improvement generated by the segment-syllable help treatment (see Table 5.4 and Figure 5.1). Again this is further discussed in section 6.2.3 of the Discussion chapter.

Table 5.4. Tukey homogeneous subsets for word form error change from pre-test to post-test in treated words

Study groups	N	Subsets for alpha =0.05			
		1	2	3	4
Highlighting	50	-14.7400			
Segment	50	-14.6400	-14.6400		
Syllable	50		-13.0400		
Segment-syllable	50			-7.7000	
Control	50				-.9216
Sig.		1.000	.065	1.000	1.000

Note. The table represents the mean reduction rate of vowel blindness errors from pre-test to post-test with the significance level between nonsignificant subsets.

To summarize, the findings for RQ 1.1 reveal that there is a significant change from pre-test to post-test, indicating a reduced level of errors in word form (i.e., vowel blindness). Pairwise analysis, however, reveals a significant difference between the control and the experimental groups in reducing vowel blindness errors. Moreover, the highlighting and segment help treatments are found to be the most effective help options while segment-syllable is the least effective treatment. Relative to the pre-test scores, the reduction rate[5] is around 46% (<15 items) for the segment-focused glosses and highlighting, 41% (13 items) for the syllable-focused glosses, and 24% (7.7 items) for the segment-syllable focused glosses.

5.1.3 RQ 1.2: Over the retention period (post-test to delayed post-test), do the four different help options lead to maintenance of any reduction in vowel blindness better than no help at all? Do the four options differ in their effects on maintenance of vowel blindness reduction?

Answers to this question are obtained through the analysis of the changes from post-test to delayed post-test and whether the changes differ between groups. This provides answers to RQ 1.2 by comparing how much reduction, or loss of reduction, in vowel blindness took place due to the study help options from post-test to delayed test. For this purpose, a 2x5 mixed ANOVA was performed with alpha level of .05.

The results show a significant overall change over time (when all treatments are considered together) from post-test to delayed post-test, $F(1, 245) = 181.129$, $p<.001$, with an effect size of .424 (partial eta squared). The between-subjects test further confirms that the difference between the groups (taking both times together) is also significant: $F(4, 245) = 167.757$, $p<.001$ with an effect size of .732 (partial eta squared). In addition, and most importantly, the interaction effect of group and time is found to be significant, $F(4, 245) = 25.058$, $p<.001$ with an effect size of .289 (partial eta squared). These findings suggest that the learners' word form error scores have changed significantly over time from post-test to delayed post-test in different ways between the

[5] Reduction rate is calculated by (post-test-pre-test)/32, (delayed post-test-post-test)/32, or (delayed post-test-pre-test)/32 the total number of target/treated test items.

study groups (see Table 5.5). As Figure 5.2 shows, the changes take the form of increases in errors between these two times as is common in studies that include a retention phase (e.g., Guidi, 2009; Hulstijn, 1992; Leeman, 2003).

Table 5.5. 2x5 Mixed ANOVA of word form error changes from post-test to delayed post-test in treated words

Tests of within-subjects effects						
Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Time	1491.396	1	1491.396	181.129	.000	.424
Time*group	825.300	4	206.325	25.058	.000	.289
Error (Time)	2025.533	245				
Tests of between-subjects effects						
Intercept	81551.568	1	81551.568	4278.811	.000	.947
Group	12789.400	4	3197.350	167.757	.000	.732
Error	4688.612	245				

The change in the number of word form errors is presented in Figure 5.2 which shows that just two groups maintain almost the same level of error between post-test and delayed post-test: control and segment. All the others show some loss of the reduction in error gained between the post-test and delayed post-test, of the order of five items (16%).

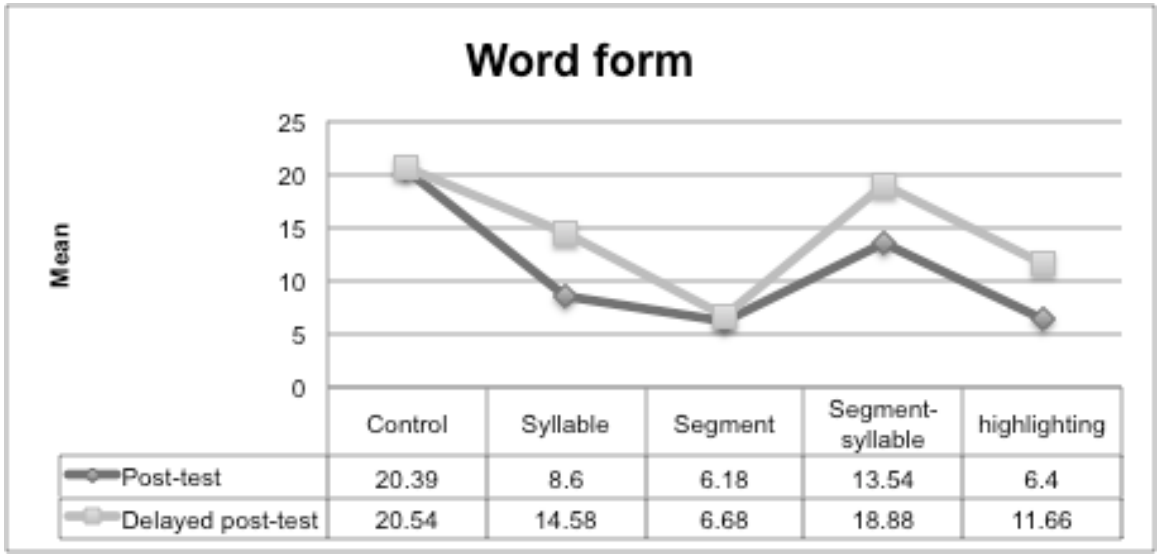


Figure 5.2. Change in vowel blindness errors from post-test to delayed post-test in treated words

Even though the control group maintains the same level of performance, it is of course doing so at a very high level of error, almost unchanged since the pre-test, around 20 items (63% error). Hence it remains the least effective in decreasing the word form errors. Most importantly, however, the segment help also retained almost the same level of word form error scores from post-test to delayed post-test. Coming on top of the achievement of almost the best error reduction in the first phase of the study, hence the segment help is clearly the best. This indicates that the segment help treatment resulted both in the best short-term decrease in vowel blindness as well as the best long-term retention of that decrease after six weeks.

Tukey multiple comparisons of the change scores (see Table 5.6. and Appendix I for Tukey's multiple comparisons) confirm these results by showing that there is no significant difference between the control and the segment group (p=.993) while there is a significant difference between them and all the other experimental groups (p<.001). Moreover, no significant difference (p=.901) is found between the highlighting, the segment-syllable, and the segment group, as presented in Table 5.6.

Table 5.6. Tukey homogeneous subsets for word form error change from post-test to delayed post-test in treated words

Study groups	N	Subsets for alpha =0.05	
		1	2
Control	50	.1569	
Segment	50	.5000	
Highlighting	50		5.2600
Segment-syllable	50		5.3400
Syllable	50		5.9800
Sig.		.993	.901

Note. The table represents the mean reduction rate of vowel blindness errors from post-test to delayed post-test with the significance level between nonsignificant subsets.

To sum up, the overall answer to RQ 1.2 is that there is a significant increase in word form error scores between post-test and delayed post-test and this change differs significantly between the groups. Among the experimental treatments, segment help was found to be the most effective in long-term retention followed at some distance by highlighting, syllable and segment-syllable help. Relative to the post-test, the loss rate

for the reduction level of vowel blindness errors (i.e., increase) is 1.5% (0.5 item) for segment-focused glosses, 16% for highlighting (5.2 item), 18% for syllable-focused glosses (5.9 item), and 16% (5.3) for the segment-syllable focused glosses. The control maintains the same high level of vowel blindness from post-test to delayed post-test.

5.1.4 RQ 1.3: Over the entire period of the study (pre-test to delayed post-test), do the four different help options yield a net reduction in vowel blindness more than no help at all? Do the four options differ in their effects on overall vowel blindness reduction?

The aim of this RQ is to examine the net effect of the treatments on the word form error scores over time, when both initial improvements and later backsliding are taken into account. To answer this question, a 2x5 mixed ANOVA was performed with pre-test vs. delayed post-test the within subject factor and the study groups as the between-subject factor.

Table 5.7. 2x5 Mixed ANOVA of word form error changes from pre-test to delayed post-test in treated words

Tests of within-subjects effects						
Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Time	5736.286	1	5736.286	783.085	.000	.761
Time*group	2949.314	4	737.329	100.656	.000	.621
Error (Time)	1802.008	245	7.325			
Tests of between-subjects effects						
Intercept	159943.379	1	159943.379	10312.978	.000	.977
Group	3354.816	4	838.704	54.079	.000	.468
Error	3815.20	245	15.509			

The results displayed in Table 5.7 reveal a significant overall change from pre-test to delayed post-test, $F(1,245) = 783.085$, $p<.001$ with an effect size of .761(partial eta squared). The between subject analysis also shows a significant difference between the study groups, $F(4,245) = 54.079$, $p<.001$ with an effect size of .468 (partial eta squared). Above all, a significant interaction effect is found between time change and groups' performance on the reduction of vowel blindness, $F(4,245) = 100.656$, $p<.001$

with an effect size of .621 (partial eta squared). These results imply that the vowel blindness errors changed in different ways between the groups over the entire period of the study from pre-test to delayed post-test.

Figure 5.3 displays the means of the groups for the change from pre-test to delayed post-test in reducing vowel blindness. The results show that the help options are effective at reducing vowel blindness overall from pre-test to delayed post-test. For instance, the control group overall reduced it only by -0.76 (under one word on average): as expected, this shows no real increase or decrease with almost the same level of word form error score from the beginning to the end of the study. Segment help is the most effective treatment by reducing word form errors overall by just over 14 out of the 32 target words (44%).

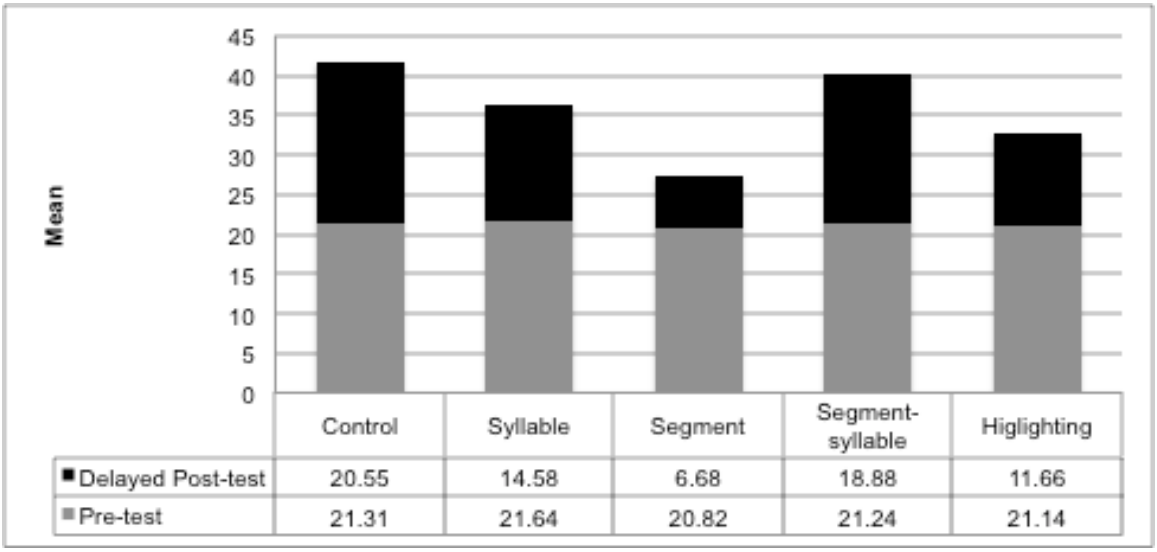


Figure 5.3. Change in vowel blindness errors for the entire study period in treated words

In order to confirm these results statistically, the Tukey post hoc test was conducted on the score differences between the pre-test and the delayed post-test. The results reveal that all the groups are significantly different from each other ($p<.001$) except for the control and segment-syllable help groups with $p=.229$ (see Table 5.8 and Appendix I for Tukey's multiple comparisons). The Tukey homogeneous subsets in Table 5.8 further show that segment help is the most effective treatment in the reduction of vowel blindness over the entire study, with a net reduction of M=14.14 items (44%)

with word form wrongly selected. Highlighting emerges as the second most effective help option treatment with a net reduction of M= 9.48 (29%) followed by syllable help with M=7.06 (22%). The segment-syllable help option is overall ineffective in reducing vowel blindness, at a level not significantly different from that of the control group.

Table 5.8. Tukey homogeneous subsets for word form error change over the entire study in treated words

Study groups	N	Subsets for alpha =0.05			
		1	2	3	4
Segment	50	-14.1400			
Highlighting	50		-9.4800		
Syllable	50			-7.0600	
Segment-syllable	50				-2.3600
Control	50				-0.7647
Sig.		1.000	1.000	1.000	0.229

Note. The table represents the mean reduction rate of vowel blindness errors from pre-test to delayed post-test with the significance level between nonsignificant subsets.

All in all, the findings for RQ 1.3 show that the net changes over the period of the study in reducing vowel blindness errors differ significantly between the types of experimental help treatment provided. The segment help had the greatest effect with a reduction of error by 14 items (44%) and the segment-syllable combination the least by 2.3 items (7.4%), indeed not doing significantly better than no help at all.

5.1.5 Summary of the findings

The analysis of findings of the first set of research questions exploring the topic of vowel blindness in relation to the experimental help options reveals the following answers:

- RQ 1.1: In the short-term from pre-test to post-test, there is a significant effect of all the help option treatments on reducing vowel blindness for the target words. However, highlighting and segment help are found be the most effective help options (around 45% reduction) with no significant difference between them while segment-syllable is the least effective treatment (24% reduction) differing significantly from all other treatment groups. Highlighting is also found to be significantly different from syllable glosses.

- RQ 1.2: In the longer term between post-test and delayed post-test, there is a significant overall loss of retention of the effects of help option treatments on reducing vowel blindness errors in target words except for the segment group. Indeed, segment help was found to be the most effective in long-term retention (1.5% loss of retention) followed by highlighting and segment-syllable help (16%). However, the main loss of retention occurred for the syllable help treatment by 18%.
- RQ 1.3: Overall, there are significant extended effects of the experimental help option treatments on vowel blindness errors in target words from pre-test to delayed post-test. All experimental groups differ significantly from each other. The segment help, however, had the greatest net effect (44% of reduction) and the segment-syllable combination the least, not differing significantly from no help at all.

5.2 Research topic 2: Transference of the effect of help options to untreated words

This research topic investigates whether the effects of the treatments on reducing vowel blindness errors transfer to untreated words. Accordingly, three research questions parallel to those in 6.1 are asked to explore this topic: 1) RQ 2.1: Over the initial intervention period (pre-test to post-test), do four different help options (in the form of input enhancement and three types of form-focused glosses) reduce vowel blindness of nontarget/untreated items more than no help at all? Do the four options differ in their effects on vowel blindness reduction of nontarget/untreated items? 2) RQ 2.2: Over the retention period (post-test to delayed post-test), do the four different help options lead to maintenance of any reduction in vowel blindness of nontarget/untreated items better than no help at all? Do the four options differ in their effects on maintenance of vowel blindness reduction of nontarget/untreated items? RQ 2.3: Over the entire period of the study (pre-test to delayed post-test), do the four different help options yield a net reduction in vowel blindness of nontarget/untreated items more than no help at all? Do the four options differ in their effects on overall vowel blindness reduction of nontarget/untreated items? To answer these research questions, form error scores for the eight distractors are analyzed to ascertain the effect of help option treatments on new words (i.e., untreated words) in the same way as we analysed the results for the targeted words above.

Before answering these research questions, an analysis of the baseline pre-test scores is provided.

5.2.1 Pre-test analysis for untreated words

The analysis of the pre-test helps to verify if the learners are all starting at the same level, especially with respect to vowel blindness. Table 5.9 provides descriptive statistics of the means and standard deviations of the control group and the four experimental groups. All groups' means are almost equal around 5 word form errors in untreated words whose meaning is known, out of 8 (63%). These results indicate that the learners enter the study with considerable familiarity with the meaning of nontarget words but still confuse the vowels in their forms at a level similar to that of the target words.

Table 5.9. Descriptive analysis of word form errors in untreated words in the pre-test

Study groups	Word form		N
	Mean	SD	
Control	5.29	1.432	50
Syllable	5.36	1.274	50
Segment	5.16	1.283	50
Segment-syllable	5.16	1.235	50
Highlighting	5.46	1.474	50
Total	5.29	1.338	250

To confirm the equal mean hypothesis, a one-way ANOVA was conducted, showing, as presented in Table 5.10, that the equal mean hypothesis is supported by a nonsignificant difference between the study groups in mean word form errors, $F(4, 245) = .468$, $p = .759$. The results confirm that all learners started at a similar level of vowel blindness with the untreated words.

Table 5.10. ANOVA results for word form errors in the pre-test in untreated words

Scoring categories	Source	Sum of Squares	Df	Mean Square	F	P
Word form	Between groups	3.378	4	.845	.468	.759
	Within groups	443.968	245	1.805		
	Total	443.347	250			

5.2.2 RQ 2.1: Over the initial intervention period (pre-test to post-test), do four different help options (in the form of input enhancement and three types of form-focused glosses) reduce vowel blindness with nontarget/untreated items more than no help at all? Do the four options differ in their effects on vowel blindness reduction with nontarget/untreated items?

To answer this research question, a 2x5 mixed ANOVA was conducted on two time points (pre-test and post-test) with the groups as between-subject factor. The dependent variable is the word form error scores which reflect the effects of the help option treatments on reducing vowel blindness.

Table 5.11. 2x5 Mixed ANOVA of word form error changes from pre-test to post-test in untreated words

Tests of within-subjects effects						
Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Time	259.852	1	259.852	401.548	.000	.620
Time*group	47.640	4	11.910	18.404	.000	.230
Error (Time)	159.193	245	0.647			
Tests of between-subjects effects						
Intercept	10471.331	1	10471.331	3156.734	.000	.928
Group	48.043	4	12.011	3.621	.007	.056
Error	816.017	245	3.317			

The findings in Table 5.11 reveal that there is a significant score change from pre-test to post-test when all groups are considered together, $F(1,245) = 401.548$, $p<.001$ with an effect size of .620 (partial eta squared). The between-subject analysis also revealed significant differences between the study groups with $F(4,245) = 3.621$,

p=.007 with an effect size of .056 (partial eta squared). The main finding, however, is the interaction between time and group which is significant with F(4,245) = 18.404, p<.001 and an effect size of .230 (partial eta squared), implying that the change in vowel blindness errors is significantly different from pre-test to post-test in a different way in different study groups.

The results prompted a follow-up analysis to ascertain where the group differences lie. A Tukey post hoc analysis was therefore carried out on the form error score differences between the post-test and the pre-test (see Appendix J for Tukey's multiple comparisons). The results show that the significant differences exist between the control group and all the other help option treatment groups (p<.001) as presented in Appendix J. The experimental treatment groups, however, are all at the same level with no significant differences between them, as shown in Table 5.12.

Table 5.12. Tukey homogeneous subsets for form error change from pre-test to post-test in untreated words

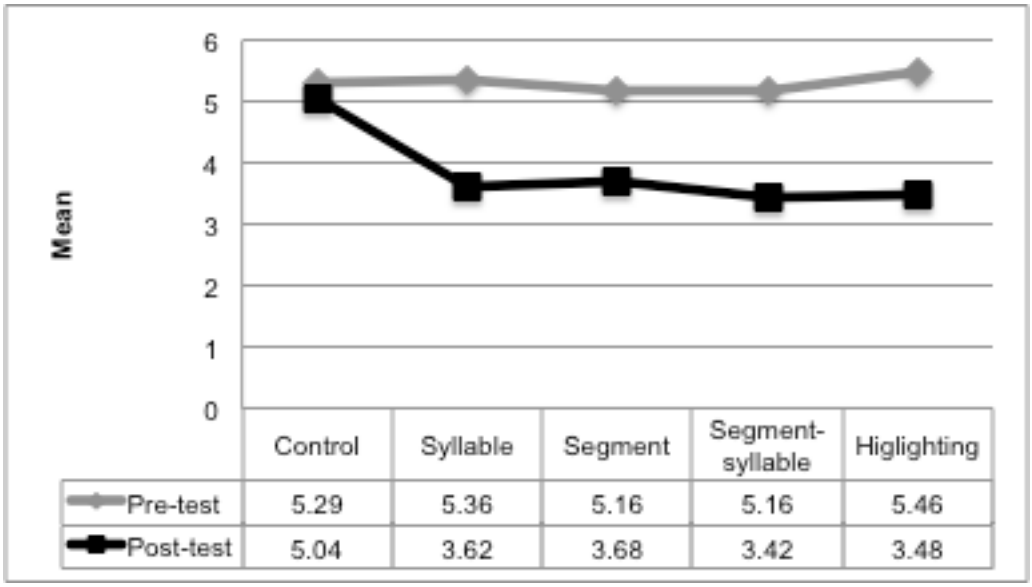
Study groups	N	Subsets for alpha =0.05	
		1	2
Highlighting	50	-1.9800	
Syllable	50	-1.7400	
Segment-syllable	50	-1.7400	
Segment	50	-1.4800	
Control	50		-.2549
Sig.		.182	1.000

Note. The table represents the mean reduction rate of vowel blindness errors from pre-test to post-test with the significance level between nonsignificant subsets.

Figure 5.4 displays the mean word form error scores in the pre-test and post-test. It shows that the help option treatments resulted in almost the same reduction of vowel blindness errors, by slightly less than two items for each treatment group (18%-24%).

In sum, results for RQ 2.1 show that the help option treatments significantly reduced vowel blindness errors more than the no help condition, even in untreated words during the initial period of intervention from pre-test to post-test. However, this

effect was at an almost equal level for all treatment groups with slightly below 2 out of the 8 words (approximately 24%).



	Control	Syllable	Segment	Segment-syllable	Higlighting
Pre-test	5.29	5.36	5.16	5.16	5.46
Post-test	5.04	3.62	3.68	3.42	3.48

Figure 5.4. Change in vowel blindness from pre-test to post-test in untreated words

5.2.3 RQ 2.2: Over the retention period (post-test to delayed post-test), do the four different help options lead to maintenance of any reduction in vowel blindness with nontarget/untreated items better than no help at all? Do the four options differ in their effects on maintenance of vowel blindness reduction with nontarget/untreated items?

To test the effects of the help option treatments on the reduction, or loss of reduction, of vowel blindness during the retention period from post-test to the delayed post-test, a 2x5 mixed ANOVA was again performed (see Table 5.13). The findings show that there is a significant overall score change from post-test to delayed post-test at $F(1,245) = 244.270$, $p<.001$ with an effect size of .498 (partial eta squared). Also, a significant difference between the study groups for both times together was found with $F(4,245) = 13.885$, $p<.001$ and an effect size of .184 (partial eta squared). Most importantly, a significant interaction effect was found of study group and time, $F(4,245) = 5.071$, $p<.001$ with an effect size of .076 (partial eta squared). These results indicate that there is a change from the post-test to the delayed post-test which operates differently between the study groups. Figure 5.5 and Table 5.14 indicate that the changes for all

groups, similar to the targeted word result, take the form of loss of reduction (i.e., increase) in word form error.

Table 5.13. 2x5 Mixed ANOVA of word form error changes from post-test to delayed post-test in untreated words

Tests of within-subjects effects						
Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Time	177.136	1	177.136	244.27	.000	.498
Time*group	14.7090	4	3.677	5.071	.001	.076
Error (Time)	178.390	245	0.725			
Tests of between-subjects effects						
Intercept	9903.999	1	9903.999	3091.419	.000	.926
Group	177.936	4	44.484	13.885	.000	.184
Error	788.112	245	3.204			

In order to ascertain where the differences lay between the groups, a Tukey post hoc analysis was conducted on the score differences from post-test to delayed post-test (see Appendix J for Tukey's multiple comparisons). The results show a significant difference between the highlighting and segment ($p<.001$) help treatments and between the highlighting and syllable ($p=.005$) treatments as shown in Appendix J. In both instances, highlighting exhibits the smallest loss (under one word). All the other pairs show nonsignificant differences between them as presented in Table 5.14 and Figure 5.5, though the segment and syllable help treatments exhibit the greatest loss of reduction, around 1.5 words (18%), which is in fact the greater part of their initial pre-post improvement. Segment-syllable glosses fall in between with a loss of reduction of approximately 1 word (12%) and by also not significantly differing from any treatment group.

Overall, there is an increase again in vowel blindness errors indicating a loss between post-test and delayed post-test of some part of the earlier reduction in vowel blindness. Highlighting, however, emerges as the most effective help option treatment in maintaining the reduction effect on vowel blindness errors from pre-test to post-test over the retention period between post-test and delayed post-test. The control group appears to maintain a similar low level of loss of earlier error reduction to that of highlighting, but

that is not especially beneficial because the control group recorded very little reduction of vowel blindness between pre-test and post-test in the first place and this is more than wiped out by the increase in error between post-test and delayed post-test.

Table 5.14. Tukey homogeneous subsets for form error change from post-test to delayed post-test in untreated words

Study groups	N	Subsets for alpha =0.05	
		1	2
Highlighting	50	.6400	
Control	50	.9804	.9804
Segment-syllable	50	1.2600	1.2600
Syllable	50		1.4800
Segment	50		1.5800
Sig.		.077	.095

Note. The table represents the mean reduction rate of vowel blindness errors from post-test to delayed post-test with the significance level between nonsignificant subsets.

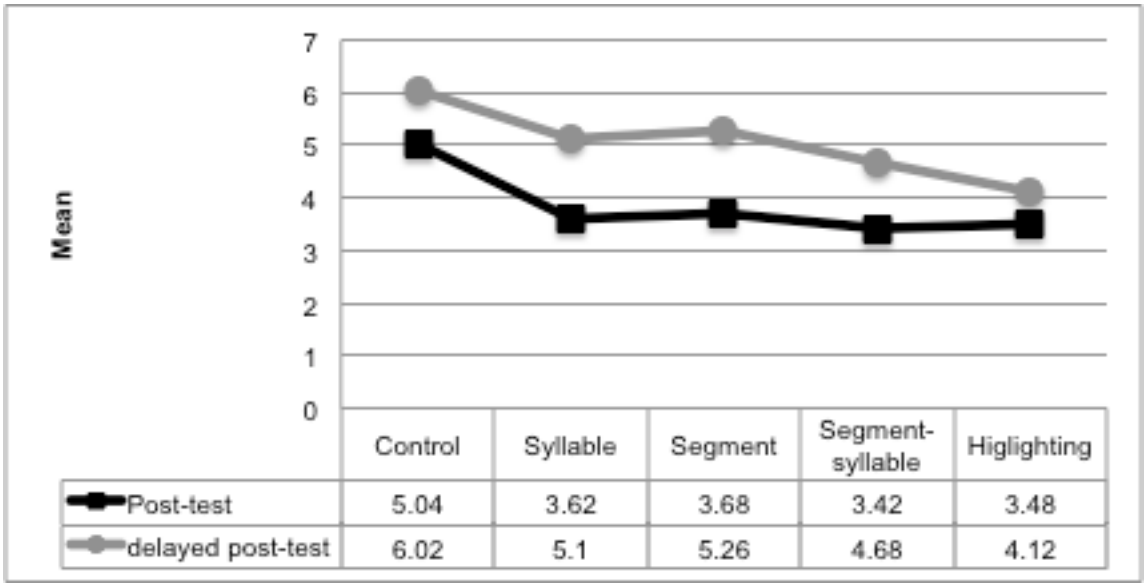


Figure 5.5. Change in vowel blindness from post-test to delayed post-test in untreated words

Taken together, the results show that there was a significant failure of retention from post-test to delayed post-test resulting in a re-increase in vowel blindness errors for the untreated words, regardless of help option. Highlighting, however, is found to be the most effective in retaining the earlier reduction effect on vowel blindness errors in

untreated words with an increase of <1 word form errors (8%) while segment help, which emerged as most effective for retention of form information about the targeted words, stands out as the least effective in this phase for the untreated words with an increase of <2 word form errors (20%).

5.2.4 RQ 2.3: Over the entire period of the study (pre-test to delayed post-test), do the four different help options yield a net reduction in vowel blindness of nontarget/untreated items more than no help at all? Do the four options differ in their effects on overall vowel blindness reduction of nontarget/untreated items?

This research question was again answered through a 2x5 mixed ANOVA between the pre-test and delayed post-test.

Table 5.15. 2x5 Mixed ANOVA of word form error changes over the entire study in untreated words

Tests of within-subjects effects						
Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Time	7.900	1	7.900	18.164	.000	.069
Time*group	58.354	4	14.589	33.544	.000	.353
Error (Time)	106.988	245	.435			
Tests of between-subjects effects						
Intercept	13372.324	1	13372.324	4259.804	.000	.945
Group	45.365	4	11.341	3.613	.007	.055
Error	772.24	245	3.139			

The findings presented in Table 5.15 reveal that there is a significant overall change in vowel blindness errors for untreated words over the entire period of the study from pre-test to delayed post-test, $F(1,245) = 18.164$, $p<.001$ with an effect size of .069 (partial eta squared) which is quite a bit less than the corresponding effect size for the target items of .761. Additionally, there is a significant difference between the groups in overall pre-test and delayed post-test scores, $F(4,245) = 3.613$, $p=.007$ with an effect size of .055 (partial eta squared). The main finding, however, is the significant interaction between time and group, $F(4,245) = 33.544$, $p<.001$ with an effect size of .353 (partial

eta squared), suggesting that the study groups change in vowel blindness errors from pre-test to delayed post-test but in different ways. Indeed, Table 5.16 and Figure 5.6 indicate that while three treatments yield the hoped for overall net reduction in vowel blindness errors in untreated words, albeit small (the syllable, segment-syllable and especially the highlighting help options), in fact two resulted in increases in error over the whole study period (the control and segment help treatments).

Table 5.16. Tukey homogeneous subsets for form error change over the entire study in untreated words

Study groups	N	Subsets for alpha =0.05			
		1	2	3	4
Highlighting	50	-1.3400			
Segment-syllable	50		-.4800		
Syllable	50		-.2600	-.2600	
Segment	50			.1000	
Control	50				.7255
Sig.		1.000	.762	.302	1.000

Note. The table represents the mean reduction rate of vowel blindness errors from pre-test to delayed post-test with the significance level between nonsignificant subsets.

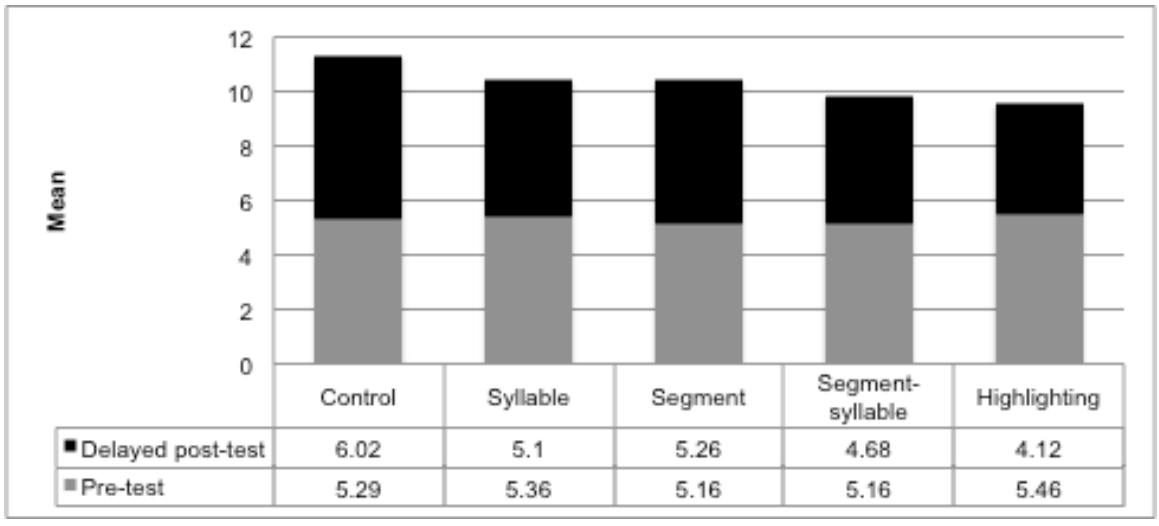


Figure 5.6. Change in vowel blindness for the entire study period in untreated words

A Tukey post hoc analysis was therefore performed to compare groups on their score differences between pre-test and delayed post-test. The findings show that

significant differences exist between the control group (with the worst error reduction) and all the other experimental groups (p<.001) (see Appendix J for Tukey's multiple comparisons). At the other end of the scale, the highlighting group with the best error reduction rate[6] of < 2 word form errors (17%) shows significant differences from all other groups (p<.001) as shown in Table 5.16 and Appendix J. The segment treatment performs the worst with a reduction of only 1 word form error (12%). However, there is no significant difference between the segment and syllable (p=.302) and the syllable and segment-syllable (p=.762) treatments as demonstrated in Table 5.16, suggesting that highlighting stands out as the most effective help treatment in terms of transferring the effect of reducing vowel blindness errors to untreated words throughout the entire period of the study.

Figure 5.6 further illustrates that the reduction in vowel blindness is mainly accounted for by the highlighting help treatment with over one word form error reduction on average (17%). The other help treatments reduce word form errors by less than half of one word form error (<6%), and in the case of segment help, a slight increase with 0.14 word form errors (2%).

Overall, some help options treatments are found to transfer their effects of reducing vowel blindness errors to untreated words, albeit at a considerably lower rate than recorded for targeted words. For instance, the net reduction effect overall of highlighting was 30% on target items but 17% on nontarget items. In contrast with the targeted items, highlighting rather than segment help emerges as most effective in transferring the effect of reducing vowel blindness throughout the study from pre-test to delayed post-test. This will be further discussed in section 6.3.3 of the Discussion chapter.

[6] Reduction rate is calculated by (post-test-pre-test)/8, (delayed post-test-post-test)/8, or (delayed post-test-pre-test)/8 the total number of nontarget/untreated test items.

5.2.5 Summary of the findings

The analysis of the second set of research questions exploring the topic of whether the effect of help option treatments extends to words not directly targeted by the help treatment reveals the following findings:

- RQ 2.1: In the short-term, during the actual intervention period from pre-test to post-test, there is a significant effect of the help option treatments on vowel blindness reduction in untreated words. However, the effect is similar for all the help options apart from no help, and below 25%. It contrasts with the corresponding result for targeted words where the effect of help ranges from 47% to 24%.
- RQ 2.2: In the longer term between post-test and delayed post-test, all treatments show a significant loss of earlier reductions in vowel blindness errors in untreated words: highlighting exhibits the least loss and segment help the most. A significant difference is found between the highlighting and segment treatments and between the highlighting and syllable treatments. Segment-syllable does not significantly differ from the remaining groups.
- RQ 2.3: Overall, from pre-test to delayed post-test there are significant transferred effects of help option treatments on vowel blindness errors in untreated words. Significant differences are found between highlighting and the remaining groups. In fact, highlighting help yields the most useful amount of error reduction overall (17%), considerably smaller than the corresponding figures for target words. Segment glosses are the least effective treatment by also significantly differing from segment-syllable and highlighting.

5.3 Research topic 3: Learners' awareness of vowel blindness and attitudes towards VALE

The third research topic focuses on the third set of research questions: 1) RQ 3.1: How satisfied were the learners with the technical design features of VALE? 2), RQ 3.2: How far did the learners perceive VALE's help options as raising their awareness of vowel blindness problems?, and 3) RQ 3.3: How far did the learners perceive VALE's help options as assisting their learning about vowels?

To answer these research questions, a combination of retrospective interview data and attitude questionnaire data are used for RQ 3.1 and RQ 3.3 (see section 4.5.1) and retrospective interview data for RQ 3.2. It is important to note that even though the participants were given the choice to speak in Arabic and English, all of them chose to

speak in Arabic. Accordingly, all the quotes presented in this chapter were translated from Arabic to English by the author of this dissertation.

The interview data were coded according to three main themes: 1) the learners' evaluation of VALE's technical design, 2) the learners' awareness of the vowel blindness problem, and 3) the learners' perception of the effectiveness of the type of help option they had received in VALE during the study. Data reflecting each theme are used to answer RQ 3.1, R.Q 3.2, and RQ 3.3, respectively.

5.3.1 RQ 3.1: How satisfied were the learners with the technical design features of VALE?

This research question concerns the learners' evaluation of VALE's technical design. The answer for this research question is mainly derived from six questions in the attitude questionnaire together with some interview data which asked about the video tutorial for using VALE, signing up, VALE design, the directory page, the timer feature, time for each reading passage, and the time for the entire study.

Figure 5.7 displays the learners' evaluation of these features in VALE. The reported scores are mean satisfaction ratings on a Likert scale from 5 to 1 where 5 refers to strongly agree and 1 to strongly disagree. The results reflect a satisfaction with the design choices of VALE by assigning the highest score to the implementation of the video tutorial (4.768) and the lowest score (2.448) to the length of phase 2.

These results are also supported by the interview data on VALE's design. One of the main themes emerging from the interview data, mainly but not exclusively from interview questions 1 and 2 (see Appendix F), is the learners' evaluation of the technical design of VALE. 370 comments were related to this theme including 319 (86.2%) positive and 51 (13.7%) negative comments. Two categories are distinguished within this theme: 1) general evaluation comments about VALE (167=45.1%), and 2) participants' favoured/disfavoured features in VALE (203=54.8%). The disfavoured features are also considered negative comments about the design of VALE.

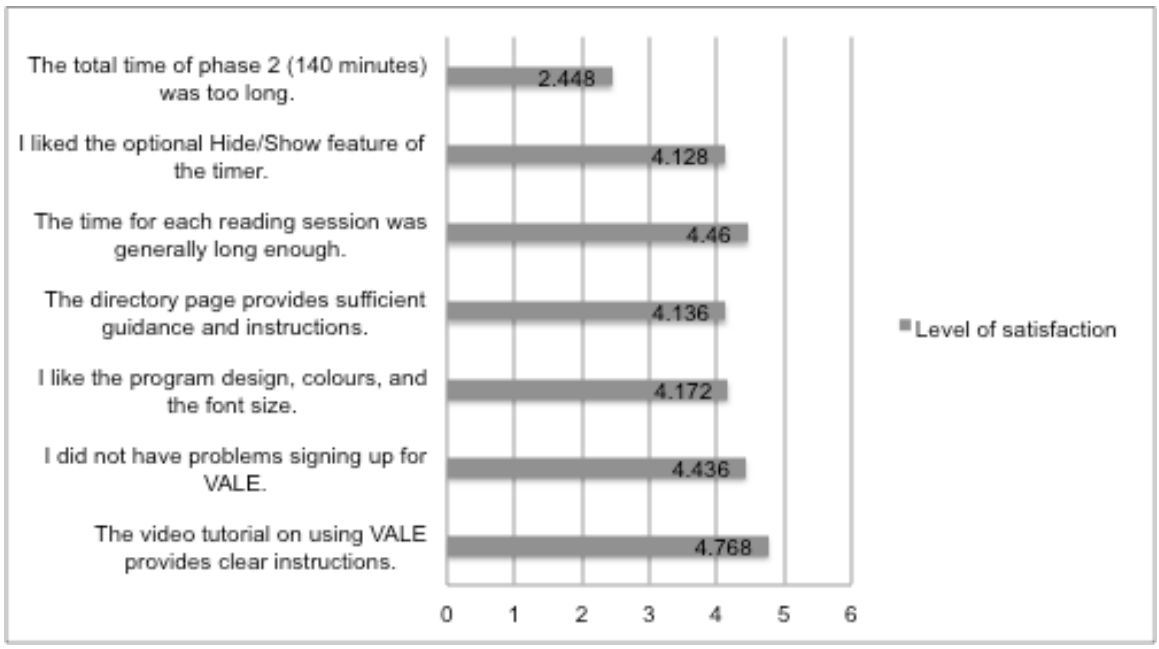


Figure 5.7. Learner evaluations of VALE's design

Generally, the learners expressed positive comments towards the technical design features of VALE, for example:

> *VALE is user-friendly. Actually, I hope I can always use it. It is very flexible and self-guided.*
>
> *The program is well-designed because it is equipped with self-learning tools such as guided stages, reading exercises, answer check, and vowel learning tools.*
>
> *VALE looks simple and easy to use. I like the design! It is non-distracting and it directs my attention to the content. I did not really feel that I am missing anything. It has perfect tools for learning especially for beginners.*
>
> *The design is very helpful because it looks simple and easy to use for beginners of English. It is even easy for learners with poor computer skills.*

The second category, favoured/disfavoured features in VALE, emerged from 203 comments about what they liked and disliked the most in VALE. Examples of the favoured features include mainly the help options (101=49.7%), the directory (24=11.8%), the content organization of the reading sessions (11=5.4%), the simple look of the program (21=10.3%), the answer check (15=7.3%), the pop-up instructions

window for every phase (14=6.8%), and the timer for tests and reading sessions (17=8.3%). For instance:

> *What I liked the most is that VALE has a very organized directory. It shows the stages, the progress, and the time for each stage.*
>
> *The feature that attracted my attention is how compact the content of the reading session is on one page with many help options such as the vowel information, the word meaning, and the reading exercises. It makes everything easily accessible.*

Among the negative points were:

> *I did not like to be locked and forced to check the help option again for a word meaning if I did not get it correct. It is frustrating especially when sometimes I just made a click mistake while I really knew the answer.*
>
> *I felt that it looks dull. More graphics and animations would give VALE a nicer look.*

5.3.1.1 RQ 3.1: Summary of the findings

Generally speaking, the technical design of VALE generated positive attitudes in the attitude questionnaire as well as positive comments from the retrospective interviews for features such as timer, directory page, time of each reading session, colours, type and the size of the font, signing up, and the video tutorial. Through the interview data more design features emerged as favoured such as the content organization of the reading sessions, the simple look of the program, the answer check, and the pop-up instructions window for every phase. However, there were a number of negative comments in the interviews (around 13.7% of the total remarks about the design of VALE). These comments are mainly about the locked out feature for the correct answer and the non-graphic look of the program. All in all, the learners' responses in the questionnaire and the interviews reflected a positive evaluation of VALE's technical design features. The next two research questions present the learners' opinions about the learnability value of the type of help options they received in VALE.

5.3.2 RQ 3.2: How far did the learners perceive VALE's help options as raising their awareness of vowel blindness problems?

This research question focuses on how learners perceive the effect of VALE's help options on awareness of vowel blindness. This RQ's answer is based on the retrospective interviews conducted with 10 participants from each of the help option groups (see section 4.4.3.1). The interview transcripts were examined and coded for evidence of awareness of vowel blindness to identify patterns, commonalities and differences among learners and to establish a connection between awareness of vowel blindness and the type of help option treatment. The interview questions that yielded the findings for this research question were questions 1, 2, 3, 5, 6, 7, and 8 to assess the general awareness of vowel blindness and interview questions 10, 11, 12, 14, 15, 16, 17, 18, and 19 to examine the connection between awareness of vowel blindness and help option treatment (see Appendix F).

There are a total of 152 comments which include general comments about awareness of vowel blindness (106=69.7%), and comments about awareness of vowel blindness based on the type of the help option treatment the study participants experienced (46=30.2%). The evidence provided in the following section is grouped into two main topics: general awareness of vowel blindness and awareness of vowel blindness based on the type of help option treatment.

5.3.2.1 Awareness of vowel blindness in general

Concerning vowel blindness in general, a total of 106 comments were received from 40 participants (69.7%). Some students just reported that their general awareness of vowels had changed overall, due to participation in one of the experimental groups:

> *The study is beneficial because now I know that vowels are affecting my English skills.*

A few reported that they had in fact been aware of the problem before the study, but the help treatment that they received helped them understand the nature of the problem better:

I have always had a problem in reading and spelling words with the right vowels. VALE training helped me understand my problem.

I always think that I have problems in reading, especially when the words look similar with letters such as e, i, o and so on. [The interviewer, you mean vowels?] Yes! The vowels are very difficult in reading. I could not choose the right word easily although I knew their meanings but it was still very confusing. After reading with vowel information, I find myself knowing the problem and how to overcome it.

Rather more responses from the participants (77 out of 106, 72.6%), however, implied a lack of awareness of the problem before the study. Some of them stated quite strongly how important they now realise this awareness to be for reading and writing:

I never guessed that the vowels are my main problem in reading and writing English.

I need to pay attention to vowels when reading and spelling words because I did not know it was a source of so much confusion in reading and spelling English before the study.

I never thought that vowels were so important for my reading as I do now after VALE training.

Interestingly, many of the comments mentioned the tests as in some way involved in raising their awareness, not just the vowel help information, despite the fact that the aim of the research was to assess the effectiveness of the help options in raising awareness, not the effect of testing on awareness. That is to say, the testing was intended to gather data for the researcher on the effects of the help options, not to be itself a source of help/awareness raising. This issue will be further discussed in section 6.4.3 of the Discussion chapter.

Some students drew attention to the part of the test that made them more aware of form problems, right from the pre-test, although the post-test really clarified the nature of the problem. Recall that the test items were in pairs with the first one testing vowel knowledge while the second tested knowledge of meaning:

I found the first parts of the tests were more confusing than the second parts because I had difficulty in choosing the right form. I

found it hard also to choose the right meaning in some cases. In the post-test, the training improved my knowledge of vowels a lot.

From the beginning of the pre-test, I felt that these were easy words but I still could not choose the right form. In the post-test, I realized that the word options were different only in vowels.

Others laid greater emphasis on the post-test as being the defining point where vowel blindness was noticed:

I realized that I have a problem with the vowels during the study and mostly in the post-test when I found myself doing better after the training than I was doing in the pre-test.

I did not notice in the pre-test that the questions' choices were different because of the vowels but in the post-test and the delayed-test I found myself more aware of my vowel problems after the study.

Finally, several students offered as a reason for their lack of earlier awareness that their teachers had never focused on vowels sufficiently. For instance:

My teachers did not teach vowels in such detail as done in the study. Therefore, I did not know that I have a problem in reading vowels.

Consequently one study participant recommended that teaching should change:

I think my teachers should tell us to focus on vowels and explain that it will help us avoid confusing similar words.

Another student, however, saw it as her own task to deal with the problem in future:

Through the study, I realized my problems with vowels. I will try to always pay attention to vowels in reading.

5.3.2.2 Awareness of vowel blindness and the individual help options

The interviews also provided 46 out of a total of 152 (30.2%) specific comments about the reported role of each help option in raising the learners' awareness of vowel blindness. These comments were then subdivided into four categories according to each help option: segment, syllable, segment-syllable, and highlighting.

Out of a total of 46 (32.6%) comments concerned the role of segment glosses in raising the learners' awareness of vowel blindness, and all were positive. For instance, one participant mentioned this very explicitly:

> *VALE is very helpful to me. It makes me realize three important things about my English study: 1) I have a problem with vowels, 2) I need to pay attention to vowels when reading, and 3) I need to read and write more than I do now.*

Another student mentioned it more as an aside:

> *The segment training is perfect for overcoming the vowel problem which I happen to know from this study.*

Others showed that beyond making them aware of the vowel problem, the help enabled them to understand the problem better:

> *Through the help tools with vowel sounds and pronunciation, I was able to understand my vowel mistakes in the pre-test.*

Interestingly, one student mentioned the value of the audio feature of the segment help in developing her understanding of the problem:

> *The help information about vowels with audio, vowel letters, and symbols increased my knowledge about the vowel problems I had in the pre-test a lot.*

The syllable-focused glosses group generated 9 out of a total of 46 (19.5%) responses. 8 out of the 9 (88.8%) comments were positive. For instance:

> *The more I go through the training [the syllable focused glosses], the more I realize my problems with vowels.*

Another student pointed out that syllable help assisted not only with noticing but also with understanding the problem.

> *It makes it easy for the English learner to read the vowels by dividing the words. The vowel training helps me notice and understand my vowel problem.*

Furthermore, the following student also stressed the role of the tests together with the syllable training in raising awareness of the vowels (cf. 5.3.2.1):

> *I realized my vowel problem through the tests and then the syllable training.*

One of the participants, however, did make a negative remark about the syllable help option, implying that she did not find it improved her vowel awareness. Notably she refers first to the rules, which were in fact stated in a more demanding metalinguistic form for the syllable help than the segment help:

> *I was not really sure why I was receiving the syllable rules and treatment. It was confusing to me.*

The third category of help options was the segment-syllable focused glosses which yielded 12 out of a total of 46 (26%) comments on awareness of the vowel blindness problem. 9 of the 12 (75%) comments were positive:

> *The training is important because it made me aware of my vowel problems.*
>
> *During the reading part, I realized my struggle with vowels.*
>
> *Yes, I noticed my vowel problem through the study. The training raised my awareness of the vowel problem.*

There were three negative remarks (25%) on the impact of the segment-syllable help option on awareness of vowel blindness. One student was in fact positive about the awareness raising value of the help, but not about its value for learning how to solve the problem, due to the amount of information presented. The latter will be discussed in detail in relation to RQ 3.2 (see section 6.4.3 of the Discussion chapter):

> *It looks like very good training and I felt that I was learning from the first two passages but after that, the information for all the words became so much to learn. However, it taught me that vowels are a big problem if I do not pay attention to them.*

The other negative comments showed that awareness was not raised by this form of vowel help. Again, the amount of information was given as the reason, which is understandable as this form of help gave more than the others, including the demanding

metalinguistic material from the syllable help but not the easier to process vowel audio link from the segment help:

> *The help box was loaded with so much information and it was hard for me to concentrate on any of it. I was not aware that vowels were the source of confusion in the tests and the help option did not help me either.*
>
> *I did not like my training [segment-syllable]! I was not aware of why I was reading a lot of information about vowel letters such o, e, and so on.*

Finally, the impact of highlighting as a form of input enhancement on raising the learners' awareness of the vowel blindness problem was also overall positive. There were 10 (21.7%) comments of which 8 (80%) were positive. As one student strongly stated:

> *Highlighting is very effective in attracting my attention to vowels. I did not know that vowels could be such a serious problem for my reading skill.*

This is understandable perhaps, given that this was the only help option which picked out the vowels in the words in the reading text itself, as against targeting them in a pop-up box which appeared after a word was clicked.

Two other students commented on the role of the tests together with highlighting in raising their awareness of vowel blindness (as seen before, cf. 5.3.2.1):

> *I noticed in the pre-test how hard it is to choose when words have the same consonants but different vowels. Through the highlighting of the vowels, I started to be more careful in paying attention to the word and more to the vowels. It makes me aware of my vowel problem.*
>
> *The yellow highlighting of the vowels helped me to look at the word carefully. I was not sure why they were highlighted at the beginning. Then, I remembered my problem in the pre-test to choose the right word based on the vowels. It made me realize that my problem was with the vowels and it helped me notice the vowels in these words.*

On the negative side, however, two students reported a sense of lack of awareness and distraction caused by the highlighting of the vowels:

The yellow highlighting is distracting and I could not figure out why these letters were highlighted.

I was not aware that the word form which was part of the tests was difficult because of the vowels and highlighting did not help me with that.

5.3.2.3 RQ 3.2: Summary of the findings

In sum, the findings from the qualitative data reveal that the study was successful in raising many learners' awareness of their vowel blindness problem. The learners also perceived the impact as generally positive for their further understanding of the vowel blindness problem. A very interesting finding, however, relates to the role of the tests in prompting noticing / raising awareness of the vowel blindness problem, in combination with the help options. With respect to the individual help options, the learners differed slightly in how beneficial they perceived the help options to be in raising awareness of their vowel problems. As summarized in Table 5.17, the segment group provided the most comments (15=33%) which are all positive, while the segment-syllable group produced 12 (26%) remarks out of which 10 are positive. The highlighting group generated 10 comments (21.7%) out of which 8 are positive. Finally, the syllable group made 9 (19.5%) comments out of which 8 were positive.

Table 5.17. Summary of the interview comments in percent

Help options	Percent of all comments	Percent of comments that were positive
Segment	33	100
Syllable	19	89
Segment-syllable	26	83
Highlighting	22	80

5.3.3 RQ 3.3: How far did the learners perceive VALE's help options as assisting their learning about vowels?

The final research question focuses on the learners' attitudes and perception of the learnability value of whichever help option they received in the study. Once again the data come partly from the attitude questionnaire (see section 4.4.3) and partly from

interviews. The data broadly divide into two categories: 1) the ease of the help options, and 2) the usefulness of the help options.

For the relative ease of use of the help information, as revealed by the questionnaire data, Table 5.18 reports the mean levels of satisfaction (on the scale where 1 indicates strong disagreement and 5 strong agreement) for attitude questionnaire items relevant to the three treatments which provided clickable glosses (see Appendix D). Of the five items that relate to ease, the segment group on average always gives the response that shows the most satisfaction, followed by the syllable and segment-syllable groups in that order.

Table 5.18. Responses to the ease of help option attitude items concerning form focused glosses, by group

VALE feature	Glosses groups	Level of satisfaction
		Mean
The information about the vowels is easy to understand.	Segment	4.46
	Syllable	4.28
	Segment-syllable	4.22
The information about the vowels is helpful.	Segment	4.48
	Syllable	4.32
	Segment-syllable	3.60
The information in the glosses seems over-loaded.	Segment	1.94
	Syllable	2.52
	Segment-Syllable	2.84
I had a hard time reading the vowel information in the glosses.	Segment	1.92
	Syllable	2.78
	Segment-syllable	3.04
The vowel information is hard to understand.	Segment	2.10
	Syllable	2.14
	Segment-syllable	2.54

Turning now to the usefulness of the help, Table 5.19 shows that all help option groups strongly agree with the statement “I do not need vowel help options” while the highlighting group is showing the highest average level of disagreement of 1.3. Moreover, the highest average of agreement is recorded for the segment group of 4.6 to

the statement “help options improved my reading skills and noticing of vowels” while the average level of agreement for all help option groups is 4.2. The item "I did not want to check vowel help options for all the words but VALE forced me to do so" attracted most agreement from the segment group with average of agreement of 2.8 while the average of the agreement for all the other groups with help glosses is 2.6.

Table 5.19. Responses to the usefulness of help option attitude items concerning form focused glosses and input enhancement, by group

VALE feature	Help groups	Level of satisfaction
		Mean
I do not need vowel help options.	Segment	1.60
	Syllable	1.90
	Segment-syllable	2.00
	Highlighting	1.30
Help options improved my reading skills and noticing of vowels.	Segment	4.60
	Syllable	4.00
	Segment-syllable	3.90
	Highlighting	4.30
I did not want to check vowel help options for all the words but VALE forced me to do so.	Segment	2.80
	Syllable	2.58
	Segment-syllable	2.62
The audio pronunciation of the word helped in learning the vowel pronunciation of the word.	Segment-syllable	4.00
The audio pronunciation of vowel sounds helped in reading vowel letters in English words.	Segment	4.30
The syllable types helped me in using vowel information in order to read English words.	Syllable	4.00
	Segment-syllable	3.92
It is easy to read by dividing the word into syllables.	Syllable	4.20
	Segment-syllable	3.90
Dividing the word into syllables helped me remember the word form better.	Syllable	3.90
	Segment-syllable	3.60
I did not understand why some letters are highlighted in yellow.	Highlighting	1.05
I noticed that vowels are highlighted for some words in the reading texts.	Highlighting	4.4

Other features of the help options that were shared only by a subset of the study participants were also generally rated high for usefulness. The learners in the segment and segment-syllable group reported that audio pronunciation of the vowels and the words helped them in learning and reading the vowels with an average of 4.3 agreement in the segment group and 4.0 agreement in the segment-syllable group. Moreover, the syllable and the segment-syllable groups reported that syllabification helped them in using vowel information to read English words, with 3.9 agreement in the segment-syllable group and 4.2 agreement in the syllable group. Also, they reported that dividing the words into syllables helped in retention, with 3.6 agreement for the segment-syllable group and 3.9 agreement for the syllable group. The highlighting group confirmed that the highlighting attracted their attention to vowels with an average of 4.4 agreement, which is one of the highest agreement means in the entire questionnaire results.

The interview data provided mostly positive comments supporting the findings from the attitude questionnaire. The comments about the learners' attitudes and perception of their group help option resulted in 574 remarks with 511 (89%) positive and 87 (15.1%) negative remarks. The following sections provide the descriptive results categorized by help option.

5.3.3.1 Input enhancement

The highlighting group reported 178 comments about highlighting and its overall effect. These comments include 169 (94.9%) positive and 9 (5%) negative comments.

Several comments reflected a positive attitude towards the ease of processing associated with highlighting, with respect to the advantage of uninterrupted and smooth reading. One student responded:

> *I like the way I can focus on vowels while reading… I would not choose any other help option group because they all have information which requires more reading while highlighting does not. It makes reading smooth and more focused.*

This learner clearly found that highlighting was the best help option due to the more limited information provided. It also shows some awareness of the fact that more information indicates more processing, or less focus. Other learners commented:

Highlighting makes it easy for me to notice the vowels while reading.

The vowels were very easy to notice and remember because of the yellow highlighting.

Highlighting makes me read the vowels carefully yet, effortlessly.

Apparently, all the learners agree on the ease of reading with the highlighting help option. Most of them refer to the advantage of noticing and retention of vowels through careful reading of vowels.

The usefulness of highlighting was also reflected in the following comment, which also interestingly reports its transferability to untreated words. This finding supports the quantitative results on the impact of highlighting as the most effective help option in transferring the vowel training effect to nontargeted items (Figure 6.16, and compare our discussion of this issue above for segment help):

Highlighting is very beneficial in attracting my attention to the highlighted vowels and reading them. I found myself more careful even in reading the non-highlighted words, especially the vowels.

Some negative comments are also reported. For example, some comments showed that highlighting could be a source of difficulty and for some learners not be useful due to creating distraction from the point of highlighting as well as from reading the text:

I wasted so much time just reading the highlighted vowels and words but not the text because I did not know why they were there.

The highlighting is distracting and confusing. I tried to ignore it but it became more and more confusing.

5.3.3.2 Segment-focused glosses

The segment group produced 159 remarks with 141 (88.6%) positive and 18 (11.3%) negative statements. Some positive comments focused especially on ease of use, reflecting the high attitude evaluations seen in Table 5.17 above. For instance, some learners explicitly expressed ease of processing due to the simplicity and ease of segment help information, for instance:

The information is very simple and easy to understand.

The vowel and word audio recordings simplified the information content of the help glosses.

In the following, the learner stressed the usefulness of the grapheme-phoneme distinction presented in the segment glosses. The comment suggests that the learner lacked the phonological knowledge about English deep orthography prior to the study.

The information in the glosses is like a self-lesson which is simple yet informative. It makes it easy for me to know the sounds of vowels and how they are written. I used to spell words based on the way they sound but I can see how it can be different.

Others commented directly on the usefulness of the various features of the segment help for learning, such as the vowel audio, word audio, and phoneme-grapheme information.

The vowel and word pronunciation are great in teaching me how to read vowels. I also like the way vowels are taught with the difference between the letter and sound representation.

Vowels are made very clear by the information provided for clickable words. I mostly like the pronunciation of vowels and words and the explanation at the bottom about the vowel letters and sounds of each word.

Again, the grapheme-phoneme distinction has appeared in these quotes as well in many other learners' quotes from the segment group. This indicates some general agreement about the usefulness of this feature as well as about the vowel and word audio feature. Indeed, descriptive statistics of frequency and percentage reveal that there were 56 (39.7%) remarks about the grapheme-phoneme distinction and 29 (20.5%) about the vowel and word audio feature out of all the 141 positive comments about segment-focused glosses.

Negative attitudes were also expressed towards segment-focused glosses. For instance, one student noted a preference for highlighting rather than segment-focused glosses clearly on grounds of ease:

I would choose the highlighting group because it is a burden to read and study extra information.

Another student pointed out that visual aids would have been more helpful than text-based segment-focused glosses. The learner referred to the need to reduce the text and add video where the information is probably best illustrated with visual-audio channels:

I did not learn much from the information in the glosses. I feel I need videos more than written glosses, especially because I am a beginner.

5.3.3.3 Syllable-focused glosses

In the syllable group, the remarks were 107 in total with 84 (78.5%) positive and 23 (21.4%) negative statements. Positive attitudes were again shown towards the ease of this kind of information, but with the admission that it took some time getting used to:

The syllables method of learning words is easy after a little bit of practice. The study provided sufficient practice through the words in each reading session.

Others pointed out the ease of processing and learning the syllable help information due to the minimal number of rules and the division of the words into smaller syllables:

Having only six rules for the syllables is so easy to learn and practice.

I find dividing the words into syllables makes learning the word much easier.

Many learners in the syllable help group pointed out the usefulness of syllable-focused glosses for retention, for instance:

I find it helpful for retention.

I can remember the word form correctly now by remembering the syllable types I learnt in the study.

Dividing the word into syllable parts helped me in remembering the spelling of words.

Even though the word *retention* and *remembering* occurred commonly in all the help option groups' comments, it most notably appeared in the syllable group's remarks. Learners produced 63 comments in total about retention across all the help groups. The syllable group produced 29 (46%) comments followed by the segment-syllable with 16 (25.3%) remarks and the highlighting group with 12 (19%) remarks.

Another learner interestingly referred to the usefulness of the presentation of the syllabification rules to the extent that it enhanced recall value:

> *I like the table with the syllable types and the rules about each syllable. I felt by the last reading session that I could predict the syllable type without reading the information provided but still I checked to make sure and also to register the information in my mind.*

An interesting comment also expressed the positive impact of syllabification on transferring vowel training to untreated words. This supports the finding of a statistically significant but small difference (4%) between the pre-test and the post-test scores for untreated words in vowel blindness reduction (see section 5.2.2):

> *Dividing the words into small parts is a strategy I always use to remember words better. The information in the glosses made it easy for me to remember these words and through using vowels I noticed that I can apply it to other words.*

Nevertheless, some negative attitudes toward syllable-focused glosses were reported, specifically about how hard it was to understand and apply the rules:

> *It was hard to follow the rules in the glosses and understand them.*

> *I need exercises to apply these rules about syllables and practice them. It is true I started to read the vowels but I was overwhelmed by this information, too.*

This clearly reflects the fact that the rules were longer and contained more metalinguistic information than the rules for the segment help, which attracted no such comments.

5.3.3.4 Segment-syllable focused glosses

The segment-syllable group generated 130 remarks with 93 (71.5%) positive and 37 (28.4%) negative comments. There were some positive comments about ease, though the following example dwells on the ease of working with syllables as it addresses the L1 habits of decomposing the word into smaller parts (e.g., root):

> *The syllable is very easy to read and it works well with my learning habits of breaking the words into smaller parts.*

Others picked out some help features in the glosses as facilitating their vowel learning, for instance:

> *The word pronunciation feature and the rules about difference between sounds and letters makes vowel learning simple and clear.*
>
> *Small syllables with only six rules are quite simple.*

Apparently, the student comment in the first quote is in favour of the segment help information in the glosses, especially for the audio feature and the grapheme-phoneme distinction. On the other hand, the second quote is in favour of the syllable help feature for its minimalist account of the vowels.

The grapheme-phoneme distinction appears also in the comments about the segment-syllable focused glosses. The segment group earlier showed 39.7% agreement with its positive comments indicating the usefulness of this help information in vowel learning. A finding affirmed again by the segment-syllable group which produced 10 (10.75%) of the 93 positive remarks from this group, implying its effectiveness in improving their vowel knowledge.

> *I liked the vowel letters and sounds as it makes it clear how one letter can represent different vowel sounds.*
>
> *The knowledge about the difference between sounds and letters in English has improved my word reading and spelling a lot.*

Other study participants reflected on the usefulness of the variety of information in the segment-syllable focused glosses for their vowel learning and reading skills. This reflects the fact that this help option contained more information than any of the others:

> *The glosses are very rich in content. I like the audios and the syllable rules. I learnt a lot through the syllable types and also through the vowel audio recordings.*
>
> *I am more confident about my vowel reading skills than before because the information in the glosses gives beneficial training on how to recognise the syllable types through vowels and I was taught about vowel sounds, letters, and pronunciation.*

In contrast to the positive remarks about the usefulness and effectiveness of the variety of information, some negative comments showed difficulty arising in part from the sheer amount of information:

> *It was crowded… very crowded with so much information. I felt distracted and I could not tell why.*
>
> *It was a great deal of information to process and at the same time read and do the exercises. It was so much.*

Due to the learners' language proficiency level, for some it represented too much of a reading burden

> *I am a beginner and I can hardly read properly and the information in the glosses was very complicated for me.*

5.3.3.5 RQ 3.3: Summary of the findings

The questionnaire findings on the ease and usefulness for learning of the help options showed the most positive responses from the segment group and the least positive from the segment-syllable group. The results from the interview data provided overall 89% positive remarks about the study help options. Highlighting produced the highest proportion of positive comments (94.9%) about learner attitudes and perception of the benefits from highlighting for their vowel learning. Most of the comments report highlighting's positive impact on vowel noticing and long-term retention. The segment group generated 88.6% positive remarks and mostly emphasized the beneficial role of

the grapheme-phoneme distinctions in the glosses and the audio pronunciation of vowels and words. For the syllable group, 78.5% of the comments were positive, especially about how syllabification accords with learners' cognitive habits of dividing words into smaller parts. The segment-syllable group reported 71.5% positive comments about how the combination of the audio feature and the syllable rules enriched knowledge of the vowels. However, 28.4% reported negative comments about how overloaded with information the glosses were. These comments confirm their responses on the attitude questionnaire about whether the glosses were too rich. Finally, some findings have demonstrated a similarity between groups. For instance, the syllable group and the segment group both single out the useful help effect on retention more than the other groups. Also, many comments in the highlighting and syllable groups remarked that there were useful transfer effects to untreated words.

5.4 Summary

Table 5.20 presents a summary of the main findings for each research question. The following chapter discusses these results based on the findings of previous research.

Table 5.20. Summary of the findings

Research question	Effect tested	Statistical procedure	Findings
RQ 1.1: Over the initial intervention period (pre-test to post-test) do four different help options (in the form of input enhancement and three types of form-focused glosses) reduce vowel blindness more than no help at all? Do the four options differ in their effects on vowel blindness reduction?	Change (decrease/increase) from pre-test to post-test for treated words	2x5 mixed ANOVA	Significant decrease in vowel blindness scores with group and time interaction of $p<.001$
	Differences between the help option groups	Tukey Post hoc	Significant differences between experimental groups and control. Highlighting and segment are the most effective treatments. Segment-syllable is the least effective treatment group

Research question	Effect tested	Statistical procedure	Findings
RQ 1.2: Over the retention period (post-test to delayed post-test) do the four different help options lead to maintenance of any reduction in vowel blindness better than no help at all? Do the four options differ in their effects on maintenance of vowel blindness reduction?	Change (decrease/increase) from post-test to delayed post-test for treated words	2x5 mixed ANOVA	Significant increase from post-test to delayed post-test with group and time interaction of $p<.001$
	Differences between the help option groups	Tukey Post Hoc	Significant difference between the experimental groups and control. The segment group is the most effective. Highlighting, syllable, and segment-syllable are yielding almost similar results.
RQ 1.3: Over the entire period of the study (pre-test to delayed post-test) do the four different help options yield a net reduction in vowel blindness more than no help at all? Do the four options differ in their effects on overall vowel blindness reduction?	Change (decrease/increase) from pre-test to delayed post-test for treated words	2x5 mixed ANOVA	Significant decrease of vowel blindness errors from pre-test to delayed post-test with an interaction between time and group of $p<.001$
	Differences between the help option groups	Tukey Post Hoc	The segment help shows the greatest net effect (44%) Segment-syllable combination is the least, not differing significantly from no help at all.
RQ 2.1: Over the initial intervention period (pre-test to post-test) do four different help options (in the form of input enhancement and three types of form-focused glosses) reduce vowel blindness of nontarget/untreated items more than no help at all? Do the four options differ in their effects on vowel blindness reduction of nontarget/untreated items?	Change (decrease/increase) from pre-test to post-test for untreated words	2x5 mixed ANOVA	Significant decrease of vowel blindness errors for untreated words with an interaction between time and group of $p<.001$
	Differences between the help option groups	Tukey Post Hoc	The effect is similar for all the help options apart from no help

Research question	Effect tested	Statistical procedure	Findings
RQ 2.2: Over the retention period (post-test to delayed post-test) do the four different help options lead to maintenance of any reduction in vowel blindness of nontarget/untreated items better than no help at all? Do the four options differ in their effects on maintenance of vowel blindness reduction of nontarget/untreated items?	Change (decrease/increase) from post-test to delayed post-test for untreated words	2x5 mixed ANOVA	Significant increase in vowel blindness errors with an interaction between group and time of $p<.001$
	Differences between the help option groups	Tukey Post Hoc	Highlighting is the most effective in retention of vowel blindness error reduction, segment is the least
RQ 2.3: Over the entire period of the study (pre-test to delayed post-test) do the four different help options yield a net reduction in vowel blindness of nontarget/untreated items more than no help at all? Do the four options differ in their effects on overall vowel blindness reduction of nontarget/untreated items?	Change (decrease/increase) from pre-test to delayed post-test for untreated words	2x5 mixed ANOVA	Significant change from pre-test to delayed post-test with interaction between time and groups of $p<.001$
	Differences between the help option groups	Tukey Post Hoc	Highlighting significantly yields the most error reduction overall (17%) and segment is the least effective treatment
RQ 3.1: How satisfied were the learners with the technical design features of VALE?	Evaluation of the technical design of VALE	Attitude questionnaire and retrospective interviews	Positive attitudes are generally found in the attitude questionnaire and positive comments from the retrospective interviews about the appearance and technical features of VALE
RQ 3.2: How far did the learners perceive the VALE help options as improving their awareness and understanding of vowel blindness problems?	Awareness about vowel blindness before and after the study	Retrospective interviews	The learners found their help option treatment as positively impacting their awareness of the vowel blindness problem. Segment provided 32.6% positive remarks followed by segment-syllable 26.2%, highlighting 21.7%, and syllable 19.5%.
RQ 3.3: How far did the learners perceive the VALE help options as assisting their learning about vowels?	Learners' perception of their learning experience through VALE's help options	Attitude questionnaire and retrospective interviews	The highest positive responses from the segment group and the least from the segment-syllable group

6 Discussion and Conclusion

This chapter summarizes and discusses the results presented in chapter 5 based on the findings of previous research into vowel blindness and help options. Three main sections are dedicated to discussing the findings for each research question (see Table 5.20 on p. 139 for all research questions): they cover vowel blindness and help options, transfer of the effect of help options to untreated words, and learners' awareness of vowel blindness and attitudes towards VALE. Next, the chapter presents the limitations of the dissertation study and suggests directions for future research. Finally, it concludes with pedagogical implications for teaching Arabic ESL/EFL learners.

6.1 Summary

The aim of this dissertation was to investigate the impact of different types of help options, mainly input enhancement and form-focused glosses, on reducing the vowel blindness of Arabic EFL learners. The literature review therefore discussed the phenomenon of vowel blindness and the hypothesis that Arabic ESL/EFL learners struggle with L1 interference when encoding and decoding words due to relying heavily on consonants and paying little attention to vowels (Hayes-Harb, 2006; Ryan & Meara, 1991). It also reviewed the help option studies focusing on input enhancement and glossing. From the literature review of these areas, a number of gaps were identified including: a scarcity of empirical research into vowel blindness, a lack of training employing help options to combat vowel blindness, a paucity of studies focusing on form-meaning connections through glossing, and a lack of research on input enhancement of small graphic units such as vowels (see section 3.5).

In order to address these gaps, a study was conducted based around a specially designed piece of online software, VALE, which addresses the vowel blindness problem

of Arabic EFL learners by incorporating training through help options in the form of input enhancement and form-focused glosses. Input enhancement was achieved typographically by highlighting the vowels in the target words in yellow. The form-focused glosses were designed to include either grapheme-phoneme or grapheme-syllable information, or a combination of both. VALE also delivered most of the data gathering instruments of the study, recorded participants' responses, and implemented the control treatment and the four experimental treatments designed to assist Arabic EFL learners in improving their English vowel reading.

Three sets of research question were formulated to address the gaps identified in the literature. The first and second sets of research questions addressed the effect of the VALE help options on reducing vowel blindness for the treated/targeted words and untreated/nontargeted words through comparing the change in vowel blindness errors between pre-test and post-test (initial effect), between post-test and delayed post-test (retention effect), and between pre-test and delayed post-test (overall effect). The third set of research questions explored the impact of VALE on raising participants' awareness of the vowel blindness problem and on their attitudes towards VALE. The results for each set of research questions are discussed and interpreted in the following sections in relation to the findings of previous research.

6.2 Research Topic 1: Effect of help options on targeted words

This research topic is concerned with the effect of the four help options in VALE on reducing vowel blindness of Arabic EFL learners in the words directly targeted by the help in VALE. The results were obtained for three research questions about the immediate effect of the help option treatments (see section 5.1.2), the delayed effect of the help option treatments (see section 5.1.3), and the overall effect of the help option treatments (see section 5.1.4). The results were found to be significant for the use of some of the study help options in reducing vowel blindness for the target words both in the short-term (post-test) as well as in the long-term (delayed post-test). The findings will be discussed in detail in the following sections and contextualized in relation to the existing literature on vowel blindness and help options.

6.2.1 Form-focused glosses

The results for target words reveal that the segment and syllable form-focused glosses provided in VALE are effective help options for reducing vowel blindness for Arabic EFL learners both in the short-term as well as the long-term. More specifically, the findings for RQ 1.1 indicate that vowel blindness, from pre-test to post-test, was reduced significantly by all three types of form-focused glosses. The post-test shows that vowel blindness was reduced relative to the pretest by around 46% of the words tested with the segment-focused glosses, 40% with the syllable-focused glosses, and 24% for the segment-syllable focused glosses. However, the finding for RQ 1.2 shows that the immediate effect did not last in the longer term for all three types of gloss help options. With syllable and segment-syllable glosses there was a significant loss of the improvement between the post-test and the delayed post-test, in around 16% of items tested. The segment help, however, was the most effective by exhibiting only a 1.5% increase in vowel blindness errors in items tested. Nevertheless, looking at the overall benefit from pre-test to delayed post-test in answer to RQ 1.3, a significant net reduction of vowel blindness was found both for segment and syllable gloss help when compared to the control group: only the segment-syllable help was found not to differ significantly from no help at all. The differences between the effects of individual types of form glosses, along with the input enhancement help, will be discussed in section 6.2.3.

These findings support the positive effect of help options on vocabulary learning in general (Al-Seghayer, 2001; Chun, 2011; Chun & Plass, 1996; Laufer & Hulstijn, 2001). With respect to glossing, the findings of the current study agree with previous research evidencing the effective impact it has on noticing, processing and learning L2 vocabulary. For instance, Chun and Plass (1996), Liou (1997), and Yanguas (2009) found that glosses, albeit with supplementary semantic rather than form information, were more beneficial for noticing and learning L2 words than no glosses.

More specifically, the beneficial effect of the form-focused glosses of this study in ameliorating participants' deficiencies in form-meaning mappings of the target words is in agreement with previous research by Rott (2003, 2005), Hulstijn (1992, 1993); Hulstijn et al. (1996). Rott (2003, 2005), for example, reported that additional lexical information in glosses of L2 words triggers establishment of robust form-meaning connections by

prompting noticing of the orthographic representation of the words in text and their meaning. Hulstijn (1992, 1993) also found that glosses with word meaning and form information were effective in establishing form-meaning connections. Nevertheless, he emphasizes that the robustness of these connections depends on the task or the design of the glosses, which need to induce a high degree of mental effort when processing unfamiliar words. None of the above studies in fact examined glosses designed to enhance knowledge only of form, using words whose meaning was familiar (vowel blindness). The current study thus provides additional evidence that at least some form-focused glosses are effective in getting learners to fill in their missing orthographic vowel knowledge and so establish more accurate form-meaning connections for target words.

With regard to the effectiveness of form-focused glosses alone on learning L2 words, Sanko (2006) conducted the only study that researched form-focused glosses, in contrast with meaning-focused glosses, through two modes of learning (incidental and intentional learning conditions, see section 3.3.1.3 for more details). Sanko's study reported no significant difference between the meaning-focused group and form-focused group when learning new words in English by Hungarian EFL learners. In fact, a significant difference was found between the two gloss groups and the control group with no glosses in the intentional learning condition although no significant difference was found between the glossing groups and the control group in the incidental learning condition. These findings suggest that form-focused glosses are as effective as meaning-focused glosses in learning new words, especially when learners are instructed to intentionally study the target words, which is in effect what the participants in the current study were required to do. This result is, however, somewhat misleading, in that Sanko in fact included a key piece of meaning information (L1 translation) in what are termed 'form focused' annotations as well as 'meaning focused' ones. The difference was in whether, in addition to that, the further information in each type of annotation/gloss related to form or to meaning. Furthermore, the test was primarily of knowledge of meaning (of words selected so as not to be previously known) rather than of form, since the word forms were supplied and participants simply had to select which one, by virtue of its meaning, fitted which gap in a text. There were no very similar word forms offered in the bank of alternatives such as were used in the present study: thus knowledge of form information (spelling) alone was not rigorously or separately tested.

Hence it is perhaps not surprising that both kinds of annotation/gloss performed equally well in preparing students for a test which effectively was of receptive knowledge of meaning.

All in all then, Sanko's study did not succeed either in supplying glosses that were purely form and not meaning focused, nor in testing knowledge of form independently from knowledge of meaning. This was a major reason why the current study, different from his, was designed precisely to address these features. From the findings of the current study therefore, for the first time, it can be said that at least some types of form-focused glosses are sufficient to trigger improvement in knowledge of form (specifically orthographic vowels), measured as a separate kind of lexical knowledge from meaning, in both post-test and delayed post-test conditions.

In Sanko's study, as in most of the lexical glossing studies (e.g., Chun & Plass, 1996; Hulstijn & Laufer, 2001; Hulstijn, 2003; Rott, 2005), target words were used whose meanings were planned to be initially unknown to learners. Now from a theoretical standpoint, VanPatten (2011) emphasizes that meaning processing precedes form processing. This could explain the effectiveness of the impact of the form-focused glosses in the current study since the meaning of the target words was already familiar. Hence, the processing effort could be more focused on the form, that is, the vowel information in the glosses. The familiarity of the participants with the target word meanings is demonstrated by the fact that, regardless of accuracy of vowel knowledge, participants on average chose the right meaning but the incorrect word form for 66.3% of items in the pre-test. With this level of familiarity, the processing load for meaning and message is lighter, permitting a high level of noticing and processing of the form of target words. VanPatten (1996, 2011) further states that processing of language features varies based on the communicative value invested in them, which makes some items more noticeable than others. This is supported by Guidi (2009) who attempted to examine noticing and learning of L2 grammatical items by providing meaning-focused glosses. The study examined the effect of meaning-focused glosses on inducing noticing and learning of grammatical L2 items such as the Spanish present perfect and impersonal SE. She reported that meaning-focused glosses with the L1 translation were not effective overall in triggering noticing and learning of the target L2 grammatical

features. Rather, glossing worked effectively in combination with other variables such as the type of the target linguistic item. For instance, a short-term glossing effect was found in present perfect rather than impersonal SE, which was explained as due to the fact that the former conveys more meaning. In this study, it is the targeted vowels which appear to lack the communicative value needed for noticing. However, due to targeting words whose meaning was already known and forcing participants to attend to glosses focused predominantly on vowels, with a variety of information on grapheme-phoneme and/or grapheme-syllable correspondences, word and vowel audios and IPA transcriptions, glossing was evidently successful in enhancing noticing and processing of vowels in the target words.

As for the effect of form-focused glosses in the longer term, the results show that there were varying degrees of loss of previous improvement between the post-test and the delayed post-test. The loss rate for the reduction level of vowel blindness errors is 1.5% for segment-focused glosses, 16% for highlighting, 18% for syllable-focused glosses, and 16% for the segment-syllable focused glosses. These differences will be further discussed in section 6.2.3. Despite this loss of initial improvement, however, significant differences were still obtained for the segment and syllable glosses when comparing the pre-test to the delayed post-test, implying that some level of improvement was retained in the long term.

Glossing studies, albeit targeting meaning rather than form, are generally found to yield significant results on retention of L2 vocabulary (Hegelheimer, 1998; Johnson & Heffernan, 2006; Laufer & Hill, 2000; Yoshii, 2006). It is a general pattern that some level of loss is found in any delayed follow-up but what is crucial is whether knowledge reverts almost to where it started or whether some significant treatment effect is still maintained. For instance, Laufer and Hill (2000) reported a high level of retention of 33% of all the words in their tests three weeks after receiving a dictionary glossing treatment. Chun and Plass (1996) also found a retention level of 24-26% of new L2 vocabulary two weeks after the glossing treatment with pictures, text, and videos. However, it is important to note that the delayed length of time for these is relatively short. At the same time, the findings from the current study support those results but with six weeks of delayed retention effect. Rott (2005) reported that multiple choice glosses (i.e., in-text

multiple L1 word meanings from which the learner is expected to choose the right one) were found to trigger significantly more retention than single translation glosses or no glosses at all. Rott explains that multiple choice glosses induce high-cognitive effort in processing resulting in stronger form-meaning connections in the lexicon. On the other hand, there are other glossing studies where gains reduce to a nonsignificant level when compared to a control group or to knowledge prior to receiving the glossing treatment. For example, Sanko's (2006) study yielded nonsignificant results in delayed tests for both intentional and incidental conditions in both the form-focused group and the meaning-focused group when compared to the control group. Guidi (2009) also found no significant long-term effect of glossing on noticing and learning Spanish grammatical items by English learners.

Even though those studies provided mostly meaning-related glosses to assist learning word form or meaning, the current study corroborates the positive retention effect of glossing found by Rott (2005), Chun and Plass (1996), and Laufer and Hill (2000). Laufer and Hill (2000) and Rott (2005) both argue that providing learners with multiple types of information in the glosses enhances retention of the target words due to the extensive amount of processing and multiple exposures to the word in text input and glosses. According to Chun and Plass (1996), processing the L2 input through multiple routes associated with different types of information leads to detailed decoding and enhances information retention at least in procedural memory.

With a slightly different emphasis, Paivio (1986) and Mayer (1997) emphasized, based on the dual-coding theory, that better memorization occurs when the input is processed through both verbal and non-verbal information channels: that is, the variety rather than the amount of processing is the key. The form-focused glosses in this study did not test that theory but, unlike glosses in any of the other studies discussed in this section apart from Sanko's (2006), the idea is reflected in the study design by employing not only additional textual illustration of grapheme-phoneme and/or grapheme-syllable correspondences so as to activate visual processing but also word and vowel audios to enhance auditory processing.

In addition to the amount of gloss information and its variety of modalities, a third feature often reported as aiding retention is saliency, highlighted by Chapelle (2005) who asserts that multiple modalities in help options activate dual cognitive processing of the input by providing two important elements for input retention: saliency and additional knowledge. In relation to saliency, Laufer and Hill (2000) argue that one of the reasons that good retention results are found in their studies is because the dictionary glosses were provided for target words which were highlighted in the text. The retention results in this dissertation agree with those results since the target words with form-focused glosses were underlined and assigned a distinctive blue font colour although primarily with the intent to make them appear clickable.

In vowel blindness research, a number of studies (Bowen, 2011; Stein, 2010; Taylor, 2008) have recommended and adopted different forms of phonics training. Unlike the current study, however, Taylor (2008) did not find a significant effect over a 16 week duration for the type of phonics training she used in her study. Nevertheless, she suggested more research be conducted on using phonics training to improve EFL learners' orthographic awareness and eventually overcome vowel blindness. Several differences in the form-focused glosses provided in Taylor's study and the current experiment might explain these conflicting results between the two studies. Firstly, Taylor's study investigated the effect of phonics training on developing phonological and orthographic awareness rather than the effect of the phonics training on vowel blindness. Secondly, the current study presented the phonics rules and information through glosses contextualized in a reading task, whereas in Taylor's study the rules were presented in a series of lessons structured for rote learning, detached from any reading and not in the form of glosses. Thirdly, the two studies employed different types of phonics teaching. The current study accounts for grapheme-phoneme and/or grapheme-syllable correspondences through a minimal number of rules whereas Taylor's study used a much more extensive set of rules to teach grapheme-phoneme correspondence patterns.

Section 6.2.3 will provide a detailed discussion of the effectiveness of each of the three types of form-focused glosses separately.

6.2.2 Input enhancement

The results for input enhancement of the vowels through yellow highlighting, (along with the target words being underlined and in blue font colour, as in the gloss treatments) show that this treatment was also effective in reducing vowel blindness over the whole study period from pre-test to delayed post-test. In the short term, from pre-test to post-test, input enhancement in the form of highlighting was found to be one of the most effective help options in the study in reducing vowel blindness by 46% (similar to segment form-focused glosses). It was found to be significantly different from the control group and all the other experimental groups except for segment-focused glosses. In the long term, however, highlighting yielded a significant re-increase in vowel blindness errors by 16% of total test items with a significant difference from segment help and no significant difference from syllable or segment-syllable help. Nevertheless, it still emerged overall as the second most effective help option in reducing vowel blindness six weeks after the initial treatment, significantly worse than segment gloss help but significantly better than the other gloss treatments and the control condition.

The input enhancement literature is divided by Leow (2009) into two types of research strands: conflated and non-conflated input enhancement research (see section 3.3.2). This dissertation adopted the non-conflated research strand design (see section 3.3.2.2) by comparing the effect of input enhancement in the form of highlighting with no enhancement at all (control) and with other types of help options (form-focused glosses).

The literature of non-conflated input enhancement research is generally consistent in yielding nonsignificant results for the use of textual enhancement when compared to a no enhancement group. However, very recently, Alsadoon and Heift (2015) conducted a study on the effect of input enhancement in the form of highlighting on vowel blindness of Arabic ESL learners. The study reported positive effects of the highlighting enhancement of the vowels in targeted words using similar tests to those used in this study. The authors concluded that input enhancement in the context of vowel blindness is beneficial for noticing vowels and improving participants' orthographic knowledge of the form of target words in general. The current study obtained similar results with respect to the effective impact of vowel enhancement on reducing vowel blindness in the short and longer term.

Obviously, the findings from both the present dissertation research and Alsadoon and Heift's (2015) study contrast with the remarkably consistent nonsignificant results usually found in the literature. Alsadoon and Heift (2015) argue that differences in their methodological design with respect to embedding the target words in short and simple sentences rather than a lengthy text, coupled with the type of linguistic form (i.e., vowels) being the target of learning, could explain why their findings are inconsistent with the literature. With respect to the length of the stimuli, the authors suggested that beginning Arabic ESL learners might have benefited from reading simple short sentences with target words that were familiar in meaning and partially in form. Thus, the participants' attention was likely to be more directed to noticing and deeply processing the target word form. Moreover, the authors emphasized that the type of linguistic form could have led to these significant results as non-conflated studies have all tested grammatical forms instead of vowels.

The current study corroborates all the arguments put forward by Alsadoon and Heift (2015) in addition to Laufer and Hulstijn (2001) who also assert that when the reading load is light, more attention is given to word-level reading. Even though the current study employed reading passages rather than short sentences, it still provided simple short reading texts with simple syntactic structures, familiar topics, and high-frequency words compared to lengthy and more complex texts used in prior non-conflated studies. Therefore, minimizing the processing load of reading allows for more time to be spent on processing the target forms and noticing them, especially if they are highlighted. In fact, the current study controlled for variables that might have otherwise added to the processing load and attentional span of the learners. For instance, 60% of items whose meaning was known showed vowel blindness errors in the pre-test thus implying that the target words were highly familiar in meaning and partially in form with confusion mostly in vowels.

With respect to the type of linguistic form targeted, the study again agrees with Alsadoon and Heift's (2015) argument that testing orthographic vowel forms rather than grammatical forms might have contributed to the significant effect of input enhancement on reducing vowel blindness. Shook (1994) attempted to test the impact of the type of the linguistic item combined with input enhancement. He reported that the enhanced

form of the present perfect in Spanish was more noticed and processed than its enhanced counterpart of relative pronouns. He relates his finding to VanPatten's (1989) input processing hypothesis which states that linguistic forms with communicative value (e.g., present perfect tense with aspectual meaning) are more likely to be processed before those without it (e.g., relative pronouns with a purely syntactic function). Even though vowels might carry less meaning as purely phonemic-orthographic features, Alsadoon and Heift (2015) provided evidence that their saliency for being noticed and processing can be increased through textual enhancement in the form of highlighting vowels in target words.

From a theoretical point of view, drawing attention to a target form with the goal for the learner to notice it is the main focus of input enhancement research. For this reason, input enhancement studies are generally designed to test the two-step prediction stated by Izumi (2002):

> First, the perceptual salience created by highlighting the input will draw the learner's attention to the highlighted forms. Second, once the first step is successful, learning of the attended form will occur based on the premise that attention is what mediates input and intake. (p. 567)

However, even though these studies claim to empirically investigate noticing, most of them, as Leow et al. (2003) pointed out, do not utilize an empirical measure for noticing; hence, the internal validity of their findings is in question. Therefore, Leow et al. (2003) urge input enhancement studies to use concurrent data elicitation to provide evidence of the noticing process rather than making claims based on the final product through pre-test and post-test measures. A number of studies attempted to use such concurrent data measures but still obtained conflicting results as to the efficacy of input enhancement on noticing and acquisition. The measures used included note-taking (Izumi, 2002), think-aloud protocols (Bowles, 2003; Leow et al., 2003), and eye-tracking (Alsadoon & Heift, 2015; Winke, 2013). Leow (2009a), for example, claims that think-aloud data provided evidence that learners noticed the target enhanced form for its meaning rather than the form itself. Yet, he suggests that the input enhancement merely induced a *low-level* awareness of the form which was not enough for the input form to become intake. Similarly, Alsadoon and Heift (2015) showed with their eye-tracking data that noticing of the enhanced target forms occurred by recording more eye-fixations for the enhanced

group than for the control group. However, apart from Alsadoon and Heift (2015), the studies that used concurrent data measures yielded nonsignificant findings for the effects of input enhancement.

The current study still mainly relies on the findings from pre-, post-, and delayed post-tests to infer the efficacy of input enhancement in the form of highlighting. However, it does not make claims about noticing but rather about the outcomes of using this enhancement on reducing vowel blindness errors. From an input processing point of view, Han et al. (2008) argue that the majority of input enhancement research employs a simultaneous processing model rather than a sequential processing one. Proponents of simultaneous processing assume that in incidental learning, noticing and processing of meaning and form occur simultaneously. On the other hand, proponents of sequential processing suppose that, through intentional learning, noticing and processing occur sequentially for the target input. Learners attend first to meaning and then to form, which was facilitated in the current study.

Han et al. (2008) argue for the adoption of a sequential processing model in input enhancement research for two main reasons: 1) it follows the temporal order of the meaning-based and form-based processing principle from the input processing hypothesis (VanPatten, 2011), and 2) cognitive and attentional resources are equally allocated in attending to meaning and form. To support this argument, Han et al. (2008) refer to the effective impact of the sequential processing design on noticing and learning target forms with familiar meaning in the studies by Doughty (1991) and Izumi (2002) where explicit instruction in the form of rule-production tasks is provided. The studies of Overstreet (1998) and Lee (2007) provide further support by reporting that attending to meaning comprehension distracted the learners from noticing and processing the target form. This argument explains the significant results obtained in this study for the use of input enhancement. The current study underscores the sequential processing design where the target words are familiar in meaning and only problematic in the vowel forms. Furthermore, as Alsadoon and Heift (2015) emphasized, all other variables were controlled so as to require participants solely to notice and process the vowels in the target words. As a result, in that study, a long-term effect was even obtained four weeks after receiving the enhancement treatment.

In the current study, a significant long-term effect of input enhancement was found after six weeks but with a significant re-increase in vowel blindness errors after the initial treatment intervention. According to Leow (2009b) and updated from examination of recent research, the delayed effect of input enhancement has been addressed in very few studies in either the conflated strand (e.g., Leeman, 2003; Lyddon, 2011; VanPatten & Cadierno, 1993b) or the non-conflated strand (e.g., Alsadoon & Heift, 2015; Bowles, 2003; J. White, 1998). The conflated research produced generally significant results whereas the non-conflated research was nonsignificant except for the study by Alsadoon and Heift (2015). A possible explanation is that input enhancement is made long-lived either by controlling the amount of variables to be processed as in the study of Alsadoon and Heift (2015) and the current study, or by adding more enhancement variables to boost its effect (e.g., feedback, production task, or instruction) as documented in the conflated research (see section 3.3.2.1).

The following section presents a discussion of the differences between the individual help options in the study pertaining to their effectiveness in reducing vowel blindness.

6.2.3 Differences in the effects of the individual help options

Another major aspect of the first set of research questions concerns the differences between each of the separate help options in their effect on reducing vowel blindness in the short term (RQ 1.1), long term (RQ 1.2), and overall (RQ1.3). In the short term, highlighting and segment help were found be the most effective help options (around 46% reduction in items showing vowel blindness in the immediate post-test) with no significant difference between them; segment-syllable help was the least effective treatment (24%). In the long term, the segment help was found to be the most effective for retention (1.5% loss of initial improvement), significantly better than all the other experimental groups which did not differ significantly from each other. Overall, all experimental groups differed significantly from each other over the entire period of the study. Segment help had the greatest net effect (44%), followed by highlighting (30%) and syllable help (22%); the segment-syllable combination was least effective (7%), not differing significantly from no help at all.

6.2.3.1 Segment-focused glosses

Based on the findings, segment-focused glosses proved the most effective help option throughout the stages of the study in reducing vowel blindness for the target words. One interpretation is that segment help focuses on vowels individually and directly and provided vowel audio help. Therefore, of all types of form-focused glosses, segment help provided the least demanding metalinguistic terms as compared to syllable and segment-syllable where syllable types and rules were presented. Further interpretations of this superior effect find a basis in previous vowel blindness research. Fender (2003) and Taylor (2008) claim that phonological awareness of English grapheme-phoneme mappings could improve vowel decoding skills of Arabic ESL/EFL learners. On theoretical grounds, Aro and Wimmer (2003), based on the orthographic depth hypothesis (see section 2.3.2), argue that the opacity of English orthography delays the acquisition of grapheme-phoneme mapping skills by ESL learners. Fender (2003) emphasizes that lack of phonological transparency of English orthography is in fact the essence of the vowel blindness problem for Arabic ESL/EFL learners in addition to their L1 vowel processing habits. He explains further, based on the findings of his study, that Arabic ESL learners transfer the expectation that letters in English will correspond more or less one-to-one with sounds. Accordingly, Fender (2003) suggests that Arabic ESL learners need regular exposure to the rules of English grapheme-phoneme mappings.

Several studies in the vowel blindness literature stress the importance of improving Arabic ESL/EFL learners' phonological awareness of the English orthography to overcome the vowel blindness problem (Fender, 2003; Nadia & Charles, 2011; Saigh & Schmitt, 2012; Stein, 2010; Taylor, 2008). However, the only study that attempted to investigate the impact of phonics training on raising Arabic EFL learners' phonological awareness of grapheme-phoneme correspondences in English is by Taylor (2008). This study did not find a significant effect for the segment phonics training (i.e., grapheme-phoneme) in the form of explicit systematic instruction although the phonological awareness measure revealed an increase from 36.8% to 44.7% in the post-test. In contrast, the current study found strong evidence that the use of grapheme-phoneme phonics training in the form of glosses was not only effective but actually the most

effective help used in the study to reduce vowel blindness errors of Arabic EFL learners in targeted words.

A number of reasons might explain these conflicting results. Firstly, the current study based on the recommendations and suggestions from previous research developed VALE to specifically address vowel blindness problems by providing different help options including phonological training targeting grapheme-phoneme phonics. On the other hand, Taylor's (2008) study employed *Get Reading*, a software program which was developed by a group of teachers in the UAE to teach phonics to EFL learners by adapting materials used for teaching L1 English children. Secondly, the current study contextualized the phonics information within a reading context by having learners consult the phonics information in glosses while they were reading. The aim was to improve decoding skills for items occurring during natural reading rather than through rote learning. Taylor's study provided the phonological phonics training explicitly in the form of a series of rule-based lessons. Lastly, the current study tested the effect of phonics training on reducing vowel blindness errors. Taylor, however, tested the training effect on the learners' knowledge of the grapheme-phoneme correspondences (i.e., phonological awareness) through the actual word spellings where the rules were instanced.

Regardless of the nonsignificant results, Taylor (2008) recommended further research to explore the potential of segment phonics training with Arabic ESL learners to improve their written vowel decoding skills. Saigh and Schmitt (2012) further found that L1 transfer of vowel processing routines in decoding and encoding English words for Arabic ESL learners, especially for short vowels, indicates a need for explicit phonological instruction on English orthography. The current study provides evidence supporting the arguments and suggestions from previous research about the potential role of segment phonics training in overcoming vowel blindness. Moreover, interview data from the participants confirmed the usefulness of learning grapheme-phoneme distinctions and rules presented in the segment-focused glosses. Participants' comments reflect their need for this kind of help information in order to assist processing English written vowels.

6.2.3.2 Input enhancement

Input enhancement was found to be as effective as the segment-focused glosses in the short-term. Also, it was found to be the second most effective help option in the long term with 16% loss of retention; however, with no significant differences from the syllable and segment-syllable help. In overall net reduction of vowel blindness, it emerged as the second most effective option with a net reduction of 30%, significantly worse than segment glosses but significantly better than the other options. It is somewhat surprising that input enhancement in the form of highlighting written vowels appears to be almost as effective as segment-focused glosses because their functionality is quite different: glossing provides extra vowel information (about their sounds) while input enhancement attracts the learners' attention to the vowels by highlighting their written form for noticing.

Examing the literature in the input enhancement area, the current study, as discussed in section 6.2.2, is inconsistent with previous research by yielding significant results for the effect of input enhancement on acquiring the target forms of this study (i.e., written vowels within words). However, the vowel blindness context of the study offers a number of explanations for the effectiveness of input enhancement. For instance, Randall and Meara (1988) emphasize that Arabic visual search skills in reading English has been shaped by their L1 processing of tri-consonantal root information. They emphasize that Arabic ESL learners need to be aided in enhancing their visual processing of English words in order to attend to vowel information that would not have been needed in their L1. The current study confirms that vowel blindness could be mitigated by enhancing the visual processing of vowels typographically.

From a theoretical point of view, Truscott and Sharwood Smith (2011) and Leeman et al. (1995) emphasize that noticing is the first step in converting input to intake. In the case of the current study, the target words were cognitively familiar to the learner in meaning, and partially in form, with the vowels being the only confusing part. For this reason, and as is evident in the post-test and the delayed post-test results, noticing was made easier and promoted subsequent processing of the target input (i.e., the written vowels) to become intake. Alsadoon and Heift (2015) also showed that input

enhancement in the form of highlighting was sufficient for noticing and intake of the vowels in the target words.

Another possible reason relates to the learners' proficiency level. The beginning EFL learners of the current study might have found highlighting less onerous since it entailed no extra reading effort (which the glosses did). In fact, this reason is supported by the interview data where participants remarked on several occasions the ease of reading and noticing vowels through highlighting.

6.2.3.3 Syllable-focused glosses

For the syllable-focused glosses, the study revealed effective results in the short-term with 24% of test words showing vowel blindness reduction. The syllable help group in fact performed significantly below the highlighting group but nonsignificantly different from the segment help group. Hence, initially syllable glossing produced an improvement that fell in between the top performance of the highlighting and segment help and the low performance of the segment-syllable help. In the long-term, syllable help resulted in a loss of more initial gain than any of the other types of help (18%); however, the difference between the experimental groups was not significant except for the segment group. Overall, the syllable help emerged significantly different from the other treatments with a 22% vowel blindness error reduction over the entire period of the study. Obviously, the syllable gloss help was not as effective as the segment-focused glosses or input enhancement in reducing vowel blindness errors. Yet, it was still effective to some extent in reducing vowel blindness for Arabic EFL learners, i.e., significantly better than segment-syllable glosses and no help at all (control).

The current study had included the syllabification variant of phonics teaching based on three assumptions (see section 4.4.2.4). Firstly, and based on research by Ryan and Meara (1991) and Hayes-Harb (2006), Arabic ESL learners are found to have a tendency to decompose words into smaller parts when reading English words. The syllabification phonics approach is hence thought to address Arab learners' cognitive habits of decoding words. This assumption is supported by Bowen (2011) who states that "[k]ey to this approach is segmentation. Breaking down longer words into single syllables provides smaller and easier spelling units and enables the learning of spelling

strategies" (p.80). Secondly, and based on research by Treiman and Kessler (2005) and Stein (2010), Arabic ESL/EFL learners need to develop knowledge of the consonantal patterns in English syllables so as to disambiguate the target vowels. Thirdly, in order to minimize the potential complexity of this approach, the syllabification approach used was minimal in the number of syllable patterns it identified (i.e., only six syllable patterns of very general types) by which the learners were provided with grapheme-syllable phonics information.

According to the findings of the current study, these expectations for this approach were moderately supported. Moreover, interview data from the study participants revealed that several students commented on the usefulness of the syllabification approach in supporting their strategies of breaking down the words into smaller parts (see section 5.3.3.3). While these findings, especially from interview data, confirm that Arabic EFL learners have a tendency to decompose words into smaller parts in a similar fashion to their L1 decoding habits, the learners' proficiency level as beginners might have made the reading and understanding of the syllable-focused glosses a bit of an effort.

Interestingly, four out of ten participants interviewed also indicated the benefit of the syllable-focused glosses for retention (see section 5.3.3.3). The learners perceived the effort that needed to be used when processing the syllable-focused information as making it long-lasting. This perception finds a basis in Hulstijn and Laufer's (2001) cognitive involvement load hypothesis which correlates high-mental effort with deep processing and retention of L2 input. The quantitative data, however, revealed a failure in fact to retain earlier gains, by 18% for the syllable group. One explanation is that the learners' perception might be true for retention, as more processing effort was required when reading the syllable-focused glosses. At the same time, a limitation on retention may have occurred due to the demand on working memory because the sheer quantity and quality of syllable information contained in the glosses made it harder to process than the segment-focused glosses or highlighting.

6.2.3.4 Segment-syllable focused glosses

The least effective help option in the study was the segment-syllable focused glosses which were found to reduce vowel blindness errors by 24% of the total test items in the short-term, but with a loss of 17% in the long-term. In fact, the overall reduction from pre-test to delayed post-test was nonsignificantly different from that of the control treatment, and showed an overall improvement of only 7% of items. The combined segment-syllable focused glosses were provided based on the assumption that this blended approach might address both deficiency in phonological orthographic awareness (grapheme-phoneme mappings) as well as syllabic orthographic awareness (grapheme-syllable mappings). However, the results showed that the net reduction of vowel blindness errors was nonsignficant when compared to no help at all, indicating that participants might have been overwhelmed by the amount of information provided in the glosses, resulting in a lack of beneficial effect. This interpretation is supported by several negative statements from the interview data about this type of help option (see section 5.3.3.4). Some participants stated that the amount of information in the glosses was too much for them to process and understand. Thus any benefits from the effort involved for retention, such as those described for syllable glosses, had little scope to operate. If little information is understood from a gloss in the first place, there is less benefit in the fact that the effort of gaining that understanding promotes better memory for what was understood. As will be discussed in section 6.3, however, some information must have been gleaned by participants from these glosses, since that assumption is needed to help explain the relatively better performance of segment-syllable help on untreated words.

In sum, the positive findings from the segment-focused glosses support the common assumption in the vowel blindness literature that awareness of the irregular English grapheme-phoneme mappings is mostly what Arabic EFL learners need to notice and learn in order to overcome their problems with vowel blindness (Abu-Rabia, 2002; Nadia & Charles, 2011; Saigh & Schmitt, 2012; Stein, 2010; Taylor, 2008). The similar effect of segment-focused glosses and highlighting suggests that improving awareness of English orthography, whether explicitly through supplementing segment phonics glosses or visually through highlighting the vowels for noticing and probably

further cognitive processing of vowel grapheme-phoneme associations, is beneficial in reducing the vowel blindness problems experienced by Arabic EFL learners.

6.3 Research topic 2: Transference of the effect of help options to untreated words

This research topic is concerned with examining whether the effects of the study help options in reducing vowel blindness transfer to untreated words. The second set of research questions comprised: RQ 2.1 about the initial effect of help options on improvement between pre-test and post-test (see section 5.2.2), RQ 2.2 about the retention of help option effects between post-test and delayed post-test (see section 5.2.3), and RQ 2.3 about the overall effect of help options between pre-test and delayed post-test (see section 5.2.4). The reductions in vowel blindness were overall significant in the short term with all help treatments showing reductions in vowel blindness between 19% and 25% of the words tested, yet, with considerable fall back in the longer term. The overall analysis of change between pre-test and delayed post-test, however, still reveals significant though small transfer effects of three of the help options on reducing vowel blindness errors between 3% and 17% of words tested: only segment gloss help fails to achieve any overall improvement (vowel blindness worsens by 1%). These findings will be discussed again in three sections based on the type of help options and the differences among them: form-focused glosses, input enhancement, and differences in the effects of individual help options.

6.3.1 Form-focused glosses

In the short-term, there was a significant transfer effect of all three form-focused glossing treatments on the untreated words, reducing vowel blindness errors by 18-21% of the words tested. The reduction was smaller than for the targeted words, however, where the effect ranges from 24%-46%. Nevertheless, it is an important finding, and one neglected by other studies, i.e., that some level of transfer of learning to nontargeted items can be detected. In the long-term, however, there was a significant loss in the transferred effect of form-focused glossing treatments. Only two types of form-focused glosses (segment-syllable and syllable) showed an overall significant effect in reducing

vowel blindness errors from pre-test to delayed post-test (by 3% and 6% of words, respectively). These results do, however, suggest that during the intervention period, the learners were able to benefit from the form-focused help options and even apply the information in the glosses to nontargeted words. However, this transfer effect was short-lived.

The CALL literature, including help option research, has rarely considered or tested the impact of word-focused training on words not specifically targeted by the training and, as Amonette (2001) states:

> This is a question which SLA researchers do not appear to have formally asked; however, the question of whether "transfer" exists has been one of the main research foci in the field of psychology for a century. (p.9)

Several researchers, however, have emphasized the importance of generalizing the training effects to other similar language activities. For instance, Hubbard (2004) proposes five principles for training through a CALL program where the last principle implies generalizing the strategies learned to other CALL activities. Within the glossing literature, O'Bryan (2008) attempted to assess whether learner training would have a significant effect on gloss use by following Hubbard's five step approach. However, due to time constraints, O'Bryan did not in the end pursue the last principle by testing the transfer of the training effect to other untreated words. Rott (2003) also stresses the importance of investigating the subsequent effect of glosses on the processing of text or encounters with other new words. O'Bryan (2008) and Rott (2003) both recommend gathering qualitative introspective data to obtain insights about the learners' processes in exploiting glossing when confronting new unknown words. The current study does not account for the processes of how and when the learners transfer the glossing effect to other words but it sheds some light on whether such transfer occurs and how effective it is through the eight distractors in the study tests which were used to measure the transfer effect of the help option treatments.

Amonette (2001) and Koda (2005) have pointed out that research in psychology and education could provide insights into the transfer or the generalization of training effects to words not specifically included in the training. Since the glosses in this study implement segment and/or syllable phonics approaches, research in phonics training is

deemed relevant. Several studies were found in the area of phonics training for children in their L1 and L2 testing the transfer effect of phonics training to untreated words. For instance, Thaler, Ebner, Wimmer, and Landerl (2004) reported that a word recognition computer program with phonological (i.e., letter pronunciation) and segment phonics training helped dysfluent Austrian child readers to improve their orthographic decoding of the treated words and, at a low level, the untreated words. They noted that even though the transfer of orthographic training effect was significant, it was very small in real terms.

Heikkilä, Aro, Närhi, Westerholm, and Ahonen (2013) reported that with syllable-focused phonics training the transfer effect to items not targeted by training was very low compared to the effect on trained target items. The study reported that the transfer effect is more evident with larger and less frequent syllable units. On the other hand, Lemoine, Levy, and Hutchinson (1993) did not find any transfer effect of their phonics rhyme segment training to untreated words. Overall, the current study is consistent with the results from Thaler et al. (2004) and Heikkilä et al. (2013) in evidencing a similar significant but low-level transfer effect of the segment and/or syllable phonics training in the form of glosses to untreated words.

A possible explanation for the low level of transfer effect as well as its short-lived duration in the current study relates to the limited number of untreated test items. The study included only eight untreated words (distractors) as compared to 32 targeted words, mainly to try to keep the study more manageable in terms of its duration with a fairly large sample size. The limited range of types of words necessarily included in those eight may therefore have failed to allow participants to exhibit the full extent of their transference. Moreover, the treatment period in which the learners received all the segment and/or syllable help options was very short, lasting less than two hours. Therefore, learners' attention must have been more focused on the targeted words with the glossing treatment than on other words with no treatment. The fact that a significant short-term transfer effect of up to 22% of nontarget words was still found for the gloss help options in the study, however, suggests that learners were not only studying the information in the glosses but also in some way generalizing this new knowledge to nontargeted words. Nevertheless, it is important to note that even though the results are statistically significant, they are small (less than 2 words) when compared to those of the

targeted words. According to the noticing hypothesis of Schmidt (1995), these findings might indicate conscious awareness of the targeted input leading to further mental processing of the input creating a general principle, rule, and pattern. Nevertheless, the allotted time may not have been sufficient for processing and internalization of the phonics rules into long-term memory, thus leading to the considerable drop-off in performance in the delayed post-test. Thaler et al. (2004) emphasize that a longer period of training would lead to a stronger transfer effect of phonics training.

6.3.2 Input enhancement

In the short-term, input enhancement help was found to yield a significant transfer effect of 24% reduction in vowel blindness, yet, at this point there was no significant difference between all the experimental groups. In the long term, input enhancement was found to be the most effective treatment for retaining the reduction in vowel blindness errors albeit with a drop in vowel blindness reduction effect from 25% of tested words to 17%. Overall, however, this performance is significantly better than that with any of the gloss help options, where the best net vowel blindness reduction effect (with segment-syllable glosses) was only on 6% of words tested. Once again, SLA research has not directed attention to the transfer of input enhancement to untreated words. In fact, input enhancement has been found to be consistently nonsignificant on targeted items as discussed earlier (see section 6.2.2). Consequently, it would seem to be unfitting to run an analysis on the transfer of input enhancement effect on untreated words when it is nonsignificant for treated words in the first place. Yet, in the vowel blindness context, Alsadoon and Heift (2015) provided evidence for the beneficial effect of input enhancement in the form of highlighting for reducing vowel blindness errors by Arabic ESL learners in treated words. However, their study did not examine the effect of input enhancement on untreated words.

In the conflated research strand on input enhancement, several researchers studied the impact of input enhancement in the learners' output, which, if it occurs, could be seen as a form of transfer, of receptive knowledge to productive. Schmidt (1995) emphasizes that learning occurs when the learner is not only consciously aware of the input rule, principle, and patterns but further able to produce the target structure in their

output. Therefore, the learners' output provides some insight into their ability to generalize and transfer the effect of input enhancement of the target forms into their own production. Due to the lack of empirical evidence in the literature about the transfer effect of enhancement, the current study therefore considered input enhancement studies which document an impact through the study of learners' output. For instance, Shook (1994) reported a significant effect of input enhancement on the immediate production and recognition of the target forms which were relative pronouns and present perfect forms in Spanish. Lee (2007) also reported a significant effect of input enhancement on Koreans' use of the passive voice in English, in a study which included an oral production task through discussion of the experimental text after reading it as well as an error form correction task. Both tasks showed an effect on performance with the target features in a production task (oral discussion) as well as a receptive task (error correction), thus showing transfer, since the original learning was purely receptive through input enhancement. It must be noted, however, that Leow (2009a) and Izumi (2002) assert that the design of these studies utilized learners' output as another source of learning, beside input enhancement, to direct more attention to the form rather than as a measure of transfer or generalization of the effect of the input enhancement to the use of the target form in their own production.

Accordingly, the current study contributes to the input enhancement literature in the sense that it extends Alsadoon and Heift (2015) results on the effective impact of input enhancement on reducing vowel blindness errors for Arabic ESL learners by providing evidence of the wider transfer of this effect during the treatment and even after six weeks from the post-test to delayed post-test. In sum, it reveals that input enhancement in the form of highlighting is the most effective help option in extending its effect to reducing vowel blindness errors of untreated words.

6.3.3 Differences in the effects of individual help options

In the short term, there was no significant difference between any of the experimental groups. All the help options produced quite a similar level of vowel blindness reduction between 18% and 24% of the words tested. In the longer term, highlighting was found to be the most effective in retaining the level of reduction of vowel

blindness errors in untreated words followed by segment-syllable help with no significant difference between the two treatments. Overall, highlighting emerged as clearly the most effective treatment over the entire period of the study with a significant difference from all the other experimental groups and a beneficial effect on 17% of words tested. Segment-syllable and syllable help came next with no significant difference between them while segment gloss help was the last with a significant difference from segment-syllable help but not from syllable help.

However, the results for the transferred effect of help options need to be treated with caution for two reasons. Firstly, the level of transfer was significant but small when compared to the effect size of the targeted words. For instance, the effect size for the overall reduction of vowel blindness from pre-test to delayed post-test is .621 (partial eta squared) for the targeted words and .353 (partial eta squared) for the nontargeted words. Secondly, the number of untreated words was small at eight words only, hence, differences between the effects of the help option treatments were quite small with only one or less than one test item difference.

6.3.3.1 Input enhancement

The results for the transfer effect of input enhancement on reducing vowel blindness for untreated words reveal that highlighting ranks first by performing significantly better overall than all other help options. Furthermore, and as described in 6.2, highlighting was also the second most effective help option for targeted words, thus making it the best performing option for the study considered as a whole. This result prompts an examination of the nature of input enhancement.

There are several key differences between highlighting and form-focused glosses and these might help explain the superior effect of highlighting, especially for untreated words. These relate to: 1) transferability and generalizability, 2) saliency, and 3) the learners' proficiency level.

Firstly, the effect of highlighting is less target item-specific than that of glosses. In fact, what is visually processed in the highlighting condition is not vowel-specific information. Instead, it draws attention to vowels in general rather than giving information

about them which makes it immediately applicable to all vowels without the need to retain specific information. In contrast, segment help, for instance, provided help on specific vowels, both auditory and with letter-sound correspondence rules, which not only had to be remembered but necessarily applied to the specific vowels that occurred in the target words so did not transfer so easily to other words. Robinson (1997), for instance, has documented that input enhancement with prior exposure to the target form results in more generalizability to novel stimuli than input with prior exposure but no enhancement. He emphasizes that, enhancement accompanied by prior exposure to form or the form rules, facilitates faster access to information retrieval about form in memory, hence leading to automaticity and generalizable knowledge of the form. The current study corroborates these findings by Robinson (1997) with regard to the transferability and generalizability of the input enhancement effect to untreated words.

Secondly, highlighting makes vowels more salient. It draws attention to vowels rather than supplying any information about them which is the case for form-focused rules or auditory input specified in the glosses. From vowel blindness research, Randall and Meara (1988) concluded that Arabic ESL learners need enhancement of their visual processing of English words in order to overcome L1 interference in their L2 decoding skills. Hayes-Harb (2006) and Saigh and Schmitt (2012) confirm that visual saliency of the vowels might possibly increase awareness of vowel blindness and hence could lead to less L1 interference in decoding English words.

Finally, the learners' proficiency level, as discussed earlier in section 6.2.2, might be more compatible with input enhancement as it is visual rather than auditory or cognitive (requiring thought about metalinguistic rules). Accordingly, it requires less effort because it involves only visual processing rather than phonics information processing. The reading time is also reduced since it does not require diversion from reading the text to perform additional reading of any information on vowels in glosses. Arguably, requiring less effort and time from the learner might indeed be more effective for help options provided to beginning learners, in particular. Several input enhancement researchers (Doughty & Williams, 1998; Lee, 2007; J. White, 1998) assert that input enhancement, particularly textual enhancement, is a less obstructive and burdensome technique than other techniques (e.g., corrective feedback or recast) to direct the

learners' attention to the target form. For this reason, highlighting may have scored best among all help options overall.

6.3.3.2 Segment-focused glosses

A rather surprising finding is that segment-focused glosses were found to be the least effective in transfer effect to untreated words in the long term and over the entire period of the study. In fact, they resulted in a slight increase in vowel blindness for untreated words overall although still performing significantly better than the control group. This contrasts with the finding for the targeted words which of course demonstrated that the segment gloss help followed by highlighting was the most successful in reducing vowel blindness errors (6.2.3).

The explanation for such a contradictory result from the targeted and nontargeted words is not immediately obvious, so must be somewhat speculative, though based on the nature of each gloss. Firstly, segment-focused glosses provide information for each type of vowel with IPA information, audio files, and phoneme-grapheme rules specific to the vowels in the target words. These vowels naturally, even in 32 words, could only represent a small subset of all the possible grapheme-phoneme correspondences that occur in English. Hence transferability of the information directly to any other words was necessarily limited. In contrast, syllable-focused glosses presented just six very general syllable types where each type appeared several times, and these types appeared in the untreated words (i.e., closed and open), leading to a better generalizability effect than the segment-focused glosses. Highlighting, the most effective help option with respect to transfer, as described in 6.3.3.1, of course had the advantage for transferability of providing no information specific to the vowels in any particular words.

Secondly, again in comparison with highlighting, it is important to note the different degree of explicitness of each help option. Segment help (as that in the other glosses) was highly explicit with direct rules and information about letter-to-sound correspondences whereas highlighting is less explicit with only visual enhancement of the vowels. Reber (1993) argues that less explicit learning involves item-specific memory for form noticed during training and hence it is dependent on memory of other similar occurrences. On the other hand, explicit learning is rule-based and memory-

dependent on rule information rather than on the encounter of other similar forms or instances. For our participants of relatively low English proficiency, arguably it was easier to remember and so transfer the highlighting of vowel letters than rules in English about the values of the letters. Finally, and as discussed in section 6.2.3.2, Robinson (1997) found empirically that implicit learning with enhancement is superior even to explicit rule-instruction with respect to the generalizability to new examples thereby explaining the superiority of highlighting in untreated words.

6.3.3.3 Segment-syllable focused and syllable-focused glosses

The results further show that segment-syllable focused glosses were the second most effective in retaining reduction of vowel blindness errors in untreated words followed by syllable-focused glosses. Segment-syllable help, with an overall reduction on 6% of words, yielded a slightly better performance than the syllable help (reduction 3%) though with no significant difference between the two types of help. These findings are consistent with the learners' claim that syllabification helped in retention of the target vowel forms and words (see section 5.3.3.3 and section 5.3.3.4), since retention of information is a prerequisite for its transfer to new forms. A possible interpretation then is that the learners were aware of their ability to extend the syllable-based rules to other untreated words. Another possible interpretation is that the segment-syllable and syllable-focused glosses incurred high mental effort in processing the syllable information or the combination of syllable and segment information, leading to greater retention and transfer of that effect. This interpretation finds support in the load hypothesis of Hulstijn (1992), Hulstijn and Laufer (2001), and Laufer and Hulstijn (2001) which posits that retention occurs when the learning task induces high-mental effort and processing.

Yet another explanation could be that this finding arose because segment-syllable help gave the most rules, which, while making it complicated, also made it more likely to cover learning difficulties that would occur also in nontargeted words. In short, maybe transferability can be achieved either by being very simple, so generally applicable as highlighting is, or by being very complicated, covering many eventualities as segment-syllable glosses are, but not by something in between such as segment help. By contrast, as we saw in 6.2.3, where targeted words are concerned, what

seemed to be most effective was the help that addressed precisely the vowels in the targeted words, and provided information directly about them in two modes (auditory and rule based) in clear and simple ways, so not requiring an inordinate amount of time to process the help, i.e., the segment glosses. In that part of the study, the most detailed type of gloss (segment-syllable) in fact came out worst, presumably due to its complexity and difficulty to understand.

Evidence in favour of the generality of the segment-syllable help possibly comes from the fact that, of all the help options, segment-syllable glosses come closest to maintaining the same level of performance on both treated and untreated words. While the percentages of words that benefit from a help option are generally very much lower for the untreated than the targeted words, the figures for segment-syllable help drop very little: targeted words initial improvement 24%, final overall improvement 7%; untreated words initial improvement 22%, final overall improvement 6%. This pattern is what of course makes segment-syllable help appear comparatively poor as help for treated words but stronger compared with other options as help for untreated words. Such consistency, however, might be taken to suggest that whatever mechanism produces the effect of this treatment, it operates in the same way regardless of the word (directly treated or not). This would be consistent with a set of rules that have true generality across words, as those in the glosses here arguably do. This is in contrast with glosses which are more item-specific, such as the segment gloss rules which apply specifically to the word being glossed, thus rendering that kind of gloss much better help for treated than untreated words.

6.4 Research topic 3: Learners' awareness of vowel blindness and attitudes towards VALE

This research topic is concerned with the learners' attitudes towards VALE, their learning experience, and their awareness of the vowel blindness problem. Three research questions were asked about the learners' attitudes towards the technical design of VALE (RQ 3.1), learners' awareness of vowel blindness through VALE (RQ 3. 2), and learners' perception of their learning experience through VALE. The findings were overall positive for the learners' awareness of vowel blindness and for their

attitudes towards the design of VALE as well as their learning experience. The results will be discussed in detail in the following sub-sections: technical design of VALE, perception of the learning experience in VALE, and the awareness of vowel blindness.

6.4.1 Technical design of VALE

Results for the learners' attitudes towards the technical design of VALE were derived from the attitude questionnaire data as well as from the interview data. The findings revealed positive attitudes towards VALE in general and towards features such as the timer, directory page, time of each reading session, colours, type and size of the font, signing up, and the video tutorial. From the interview data, some further features emerged as having a positive impact such as the content organization of the reading sessions, the simple look of the program, the answer check, and the pop-up instructions window for every phase. However, some negative remarks were also made (13% of total comments). These concerned the locking of the correct answer in the exercises accompanying the reading, and the non-graphic look of the program.

Insights from previous research had informed the design of VALE. For instance, several other glossing studies (Hill & Laufer, 2003; Laufer & Hill, 2000; Shalmani & Sabet, 2010) employed a count-down timer to prompt the learners about the specific time duration given for the reading sessions. In VALE, the timer can be set to SHOW/HIDE to avoid any distraction for those learners who might be distracted by this feature. A directory page is also a common design feature in many CALL and glossing studies (Al-Seghayer, 2001; Hegelheimer, 1998; Sanko, 2006; Yanguas, 2009) from which the learners are guided through the program reading passages and activities. Another feature which follows previous research was the organization on the page of the reading session content presenting the reading passage to the left, the reading exercises to the right, and the glosses at the bottom on the right (Abraham, 2007; Chun & Plass, 1996; Nikolova, 2004; Sanko, 2006). For the font and page colour, black on a white background was the standard colour used in most of the glossing and input enhancement studies whereas the targeted words appeared in blue and underlined to indicate clickable words (De Ridder, 2000; Nikolova, 2004; Park & Nassif, 2013; Simard, 2009; Winke, 2013; Yoshii, 2006).

Arguably, our adoption of these insights from previous research resulted in a positive level of satisfaction towards the general technical design of VALE. However, a few negative comments were expressed about the locked-up feature of the correct answers in the tasks accompanying the reading and the non-graphic look of the program. The locked-up feature was implemented for research purposes only and was incorporated to ensure that learners really checked the glosses and, hence, the researcher could claim that the help options had really affected vowel blindness in the targeted words. However, if the pedagogical part of VALE were being used in a normal teaching context, rather than for research, then surely this feature might become very frustrating. Hence, one might consider that a version of VALE for actual pedagogical use might omit this locking or replace it with a popup like “Are you sure you want to go on to the next reading without checking all the help?”. With respect to the non-graphic look of VALE, many comments from the interview rather than the questionnaire data, by contrast, expressed a liking for the simple look of the program. Despite these positive comments, however, if VALE were being used in a regular teaching context, rather than for research, a more graphical interface that appeals especially to a younger generation of learners could certainly be pursued. Yet, the main goal for the current study was to implement a clean design that was fairly neutral and did not distract from the actual instructional content.

6.4.2 Perception of the learning experience in VALE

The learners’ perceptions of their learning experience with VALE fell into two categories: 1) the ease of help options, and 2) the usefulness of help options. Based on the attitude questionnaire, the segment help group showed the most satisfaction with the ease and usefulness of the vowel information in the glosses, followed by the syllable and segment-syllable help group. For highlighting, similarly, a high-level of satisfaction was reported with the ease and usefulness of noticing the vowels. It is important to note that questionnaire data are reflective of the targeted words as there was no help at all provided explicitly for the untreated words. The interview data, however, does contain some remarks about applying the vowel knowledge to other untreated words in the tests used in the study.

Clearly, learners' perception of their learning experience matches the quantitative findings about the effectiveness of highlighting as the top-help option for reducing vowel blindness errors in targeted and nontargeted words, followed by segment, syllable, and segment-syllable help options.

Highlighting produced the highest proportion of positive comments (94.9%) in the interview data with regard to its ease and usefulness for vowel learning, possibly reflecting the fact that, as described earlier, it is the most effective single choice for treated and untreated words taken together, although not for treated words alone. Most of the comments mentioned the ease of processing with no interruption or distraction. Furthermore, learners stressed the beneficial effect of highlighting on noticing and retention of vowels through careful reading of the vowels. Some learners also noted the usefulness of highlighting for its impact on their handling of untreated words which occurred only in the tests. These findings from the qualitative data support the findings for the effect of help options on both the targeted and nontargeted words. They also support the finding by Alsadoon and Heift (2015) that input enhancement in the form of highlighting, in the vowel blindness context, is effective for noticing, especially when familiarity of words in meaning is ensured.

The segment group made 88.6% positive comments about their perception of the treatment effect on vowel reading and learning. As stated earlier in section 6.2.3, this confirms the argument that phonological awareness of the specific English segment grapheme-phoneme mappings seems to be the best remedy for the vowel blindness problem in Arabic ESL/EFL learners, at least in the words for which this information is directly provided (Abu-Rabia, 2002; Nadia & Charles, 2011; Saigh & Schmitt, 2012; Stein, 2010; Taylor, 2008). Furthermore, the learners revealed through the interview data some agreement on the usefulness of knowledge of grapheme-phoneme correspondences for improving their wider phonological knowledge of English orthography.

The syllabification approach was thought by the researcher to simulate an Arabic ESL/EFL tendency to break down words into smaller units to facilitate decoding and encoding (Bowen, 2011; Ryan & Meara, 1991). In fact, the participants' comments agree

with this assumption as the majority reported the usefulness of syllable-focused glosses for retention and for transference of the syllabification effect to untreated words. The actual results for the transfer effect of the help options agreed with this perception of the learners by revealing that the segment-syllable and syllable help followed highlighting as the top three significantly treatments in reducing vowel blindness errors for untreated words. This is consistent with Heikkilä et al. (2013) who emphasize the impact of syllable phonics on the retention and transference of effects to untrained words. However, some of the learners did remark on the difficulty of reading and understanding the syllable rules, thus perhaps explaining the lower level of performance for the targeted words. As we have seen, syllable help was significantly less effective than segment glosses and highlighting overall with targeted words.

Finally, learners' comments on segment-syllable help again reflected that they recognized that it addresses their L1 habits of decomposing the words into smaller parts. Bowen (2011) and Stein (2010) assert that phonemic or syllabic segmentation is beneficial for Arabic ESL/EFL learners as it provides a decoding strategy similar to their cognitive habits of decomposing words into smaller units in L1. Learners also commented on the usefulness of the variety of information in the segment-syllable help option. Indeed, the feature that might explain the superior transfer of the effect of this help option to untreated words could be the multiple sources of information. Chapelle (2003) and Lin and Chen (2007) argue that multiple modes of help are generally preferred by learners because they activate two or more channels of perception. In contrast, other learners in the study expressed negative comments about the various types of information included in the glosses by indicating that it was too much information, especially for their language level. The same difficulty was expressed for understanding the rules in the syllable-focused glosses. These findings indicate that some of the learners were overwhelmed by the amount of information, which led to a decreased effect for the targeted words where percentage improvement overall was similar to that for untreated words. In a version of VALE for actual pedagogical use, it might be useful to adjust the amount of information presented according to the learners' language proficiency level. In the account above, some discrepancies are noted between what the learners stated in the interview and questionnaire and what the quantitative data revealed. This was made more complicated by the fact that the study

draws upon two sets of quantitative findings, where the help options mostly did not behave the same in both: hence it is more difficult to ascertain whether what the students report relates more to their experiences with the treated words or the untreated ones. Such uncertainties and discrepancies are not unknown in applied linguistics in general, or in CALL in particular. Fischer (2007), for example, pointed out that researchers in CALL need to be careful about student self-report data as such discrepancies are common between what they say and what they really do. Empirically, Fischer (2004) documented a negative correlation between the learners' actual use and their self-reporting data (likert-scale questionnaire). He recommended illuminating what goes on in a more holistic way with the use of multiple sources of data such as tracking the learners' performance on a CALL task through a recording system, logging chat or discussion boards, and researcher observation. However, the current study shows a fairly good match between what learners say and what the quantitative data show with few instances of mismatch as evident in the discussions of the segment-syllable and syllable help.

In sum, learners' perception of their learning experience through VALE's help options, with a few noted exceptions, supported the quantitative findings for the effect of the help options on targeted and nontargeted words. Segment and highlighting help appeared to attract the most positive comments in the interview data and the highest level of satisfaction in the attitude questionnaire followed by the syllable and segment-syllable glosses.

6.4.3 Awareness of vowel blindness

The study also investigated the impact of VALE's help options on raising learners' awareness of the vowel blindness problem. The assumption was that awareness of the problem would result in better utilization of the remedial strategies, help options in our case, to overcome the vowel blindness problem. The results from the interview data reveal that the study was in general successful in raising the learners' awareness of their vowel decoding problems in English.

The majority of responses (72%) indicated a lack of awareness of the vowel blindness problem before the study and that is in accordance with several researchers (Bowen, 2011; Stein, 2010; Taylor, 2008) who have implied that Arabic ESL/EFL learners generally lack awareness about their vowel blindness problem. On the other hand, a few comments (23%) indicated that some participants were in fact already aware of their difficulties in reading and spelling English words with the correct vowels. However, and whether or not the learners were aware of the problem, they all affirmed that VALE with the help option treatments aided them in understanding the nature of the vowel blindness problem. Hayes-Harb (2006) reported that the participants in his study expressed conscious awareness of their overreliance on consonant information over vowel information in decoding English words. Furthermore, Hayes-Harb argues that

> If these learners are consciously aware of their word identification difficulties, or if they can be taught to be aware of them, it is possible that such conscious strategies may help them process English words. (p.337)

Findings from the learners' comments about their specific help option treatments corroborate Hayes-Harb's quote. The learners generally indicated that they benefited from their help treatments in raising their awareness of the vowel blindness problem and improving their decoding of the vowels in the target words. The segment help treatment yielded the most comments (15 = 32.6%) which were all positive. The awareness-related positive comments on the other help options ranged from 26% for segment-syllable, and 21% for highlighting, to 19% for the syllable help. Overall, these findings do to some extent correlate with the quantitative findings of the study, however. The greater appreciation of segment gloss help for awareness raising agrees with the greater benefit of this option for targeted words but not for nontargeted words. Again, the positive comments on the segment-syllable help is consistent with the earlier findings about its transfer effect to untreated words, but not with its poor performance on targeted words. Once again, a discrepancy is documented between what the qualitative and quantitative data suggest. Hence, and at least to some extent, this might support Fischer (2007, 2004) who cautioned researchers not to rely on learner's self-report data to determine effectiveness (see section 6.4.2).

A very interesting finding that emerged in relation to the awareness of vowel blindness was the role of the tests in prompting noticing/raising awareness of the vowel blindness problem, in combination with the help options. The tests were intended to measure the effect of help options rather than be a source of awareness raising. Many of the learners, however, indicated that the nature of the tests, where they had to choose from three alternative forms for each word, and where each alternative had the same consonantal structure but different vowels, directed their attention to their difficulties in choosing the right form with the correct vowels. One possible reason is that the nature of the task might have highlighted the vowel blindness problem for the learners. The multiple-choice format juxtaposed word forms with different vowels which could have helped the learners to notice more clearly their confusion about the vowels even in the pre-test, before receiving any help options at all.

Based on the concept of consequential validity by Messick (1989), test validity is not only evidential (e.g., test scores) but also consequential (e.g., value implications). Chapelle (1996) illustrates the consequential validity concept based on research in CALL as "justifications for the usefulness of an assessment for its intended purposes as well as for its unintended outcomes beyond the immediate assessment event and context" (p. 56). However, Chapelle emphasizes a lack of consequential validity research on the potential effects of CALL and testing. The current study found evidence that the study tests on vowel blindness, with multiple choices for the word spelling based on changes with the vowels, were effective in drawing the learners' attention to their vowel blindness problem as stated by many of the participants. A finding that was not intentionally investigated, yet which emerged during the analysis of the interview data on awareness of vowel blindness relates to the fact that the study tests served another purpose, i.e., directing the learners' attention to their difficulties in choosing the right form with the correct vowels. This in turn suggests that, from a pedagogical point of view, a multiple-choice recognition task such as the one implemented in the study tests of the current research might be effective in the Arabic EFL classroom in its own right to raise the learners' awareness of vowel blindness.

6.5 Shortcomings of the study

In spite of the positive findings, the current study contained several limitations which may affect the generalizability of the findings. These limitations mainly relate to: 1) the number of untreated words and duration of the help option intervention, 2) the delivery mode, and 3) the sample population.

6.5.1 Number of untreated words and treatment duration

The limited number of untreated words and the fairly short duration of the treatment phase form the main shortcomings of the study. These two weaknesses, however, both relate to the feasibility of carrying out the research study with a fairly large sample of participants.

The results of the effect of VALE's help option treatments on untreated words indicated significant transfer effects of the help treatments, albeit at a very low level when compared to the targeted words. However, the study included only eight words (i.e., the test distractors) to measure the transfer effect. This small number of words definitely limited the ability to test the effect on a representative range of words as compared to the targeted words, possibly making the results less reliable and generalizable across words. However, the choice of eight words was primarily made to keep the test length manageable for a large group of learners in order to be able to administer the treatment phase of the study within a single day thus avoiding students not returning and discontinuing the study. It was clearly more important that each help option is adequately represented and thus the size of the study sample is big enough to produce robust results despite this limited number of untreated words.

The duration of the intervention itself, that is, the reading phase with help options, was relatively short with four passages of 20 minutes each. Furthermore, the fact that the main treatment of the study was performed in a single day might have affected the learners' performance and, hence, the findings of the study. Thus, it remains possible that the differences between the help treatments might have yielded different results if they had been implemented over a longer period of time (Lyddon, 2011; Thaler et al., 2004). The study design precluded any estimation of what would happen if the help

options were used repeatedly with different texts over several weeks or months, as might happen in a real teaching situation. Thus any impact of initial novelty effect cannot be assessed, which might lead to a later decline in improvement. Nor is it possible to tell whether there would be an incremental effect and that greater vowel blindness reduction would be found on repeated use of the help options over time, especially in the form of a greater transfer effect to nontarget words.

6.5.2 Web-based delivery

VALE was delivered through a web-based application in an attempt to ease the learners' access and use without the hassle of downloading the software onto individual computers and going through an installation process. Nevertheless, the internet connection in the computer labs was very slow and there were several occasions when it went down and the students needed to log in again. Fortunately, VALE autosaves the learners' progress up to where they left off and thus their work was never lost. However, a desktop application could have eliminated the frustration arising from a slow and intermittent internet connection.

6.5.3 Sampled learner population

The study was conducted with Saudi students from only one academic institution in Saudi Arabia which is Al-Imam Mohammed bin Saud University. Also, only female students were included in the study due to a lack of access to male students. Therefore, the findings apply to the population sampled and it should be understood that other Arab learners may respond to VALE and the study help treatments differently. With Arab learners, it is important to consider background factors such as Arabic diglossia, colloquial dialects, and educational backgrounds which make one Arabic population different from another (Abu-Rabia & Taha, 2004). Taylor (2008) emphasized that care must be taken when attempting to generalize findings of Arabic ESL/EFL learners to learners from different Arabic backgrounds due to major differences in their individual and educational background characteristics.

6.6 Future research

Given the findings and the limitations of the study, possible research avenues remain open for future studies. These recommendations for future research are discussed in three areas: 1) vowel blindness, 2) glossing, and 3) input enhancement.

6.6.1 Vowel blindness

Research in the area of vowel blindness is scarce with very few studies exploring the nature of the problem. With respect to the orthographic factors that might play a role, there has been little consideration of whether the vowel position within the word (e.g., initial saliency), word length, or individual letter shape might play a role. It also might be important to examine the effect of the existence of real English words differing only in a vowel change. For example, changing the vowel in *print* produces non-words such as *prant, pront, prunt, prent*. However, changing the vowel in the word *blind* results in real words such as *bland*, *blend,* and *blond*. Possibly, and given that the vowel in *print* has little importance in distinguishing the word from others, it might be less noticed by the learner. In contrast, words such as *blind* might increase confusion because the learner sees other words with the same consonant structure but a different vowel.

A further research topic of interest relates to the difficulty in vowel perception that contributes to vowel blindness. Is it primarily a visual or an auditory problem, or both? Bowen (2011) implies that it is the combined effect of vowel blindness and vowel deafness for Arabic ESL learners that leads to the confusion in processing English vowels. Thus, future research might explore the interdependence of the auditory and visual perception of vowels. For example, little is known about the types of vowels that cause greater vowel blindness problems for Arabic ESL/EFL learners. Saigh and Schmitt (2012), for instance, examined vowel blindness with regard to the quantity of vowels and reported that short vowels cause more difficulty for Arabic ESL learners than long vowels. However, the question remains what short vowels are more difficult: for instance, short vowels that are phonetically different from or similar to those in the student's L1?

Furthermore, future research might also examine the effects of language proficiency on vowel blindness. For instance, Randall and Meara (1988) showed that even advanced learners still exhibit vowel blindness errors in their reading. An important issue further relates to how far vowel blindness errors when reading match those in production, i.e. spelling (Alsadoon and Heift, 2015). Finally, future work might also explore a wider range of instructional CALL strategies to remedy the vowel processing problems of Arabic ESL/EFL learners. For instance, future studies might examine the effectiveness of individualized feedback on the learners' errors as this might enhance their noticing of the vowels.

6.6.2 Glossing

Future research in glossing might explore gloss-based instruction of other target constructions, for instance grammatical forms, where form-focused glosses are instructional, yet delivered within a contextualized learning setting, such as attached to a reading text. In fact, most of the research in the area of glossing is targeted towards the acquisition of the meaning of L2 vocabulary as opposed to grammar. Such studies could then explore further the potential of multimedia glosses in the form of video, animation, or pictorial information, especially for beginners.

In addition, the current study focused on the product (i.e., outcomes) rather than the process of using glosses. Therefore, the validity of these findings could be further confirmed/refuted if concurrent data gathering methods were also applied. Researchers can then further explore learners' utilization of form-focused glosses during the treatment. For example, eye-tracking can be employed to track the learners' eye movement and fixations if and while reading the glosses. Think-aloud protocols were tried in the pilot study in order to explore the learners' thoughts while noticing and processing the vowels in target words and the form-focused information of the glosses. However, the attempt was not successful as the learners hardly spoke and later expressed that they were unsure what to talk about. Future research might re-try this method with prior training in think-aloud reporting before the study.

Furthermore, in order to address the vowel blindness problem, the current study created glosses which adopted phonics approaches based on segments (grapheme-phoneme) and/or syllables (grapheme-syllable). Several vowel blindness researchers (Bowen, 2011; Hayes-Harb, 2006; Saigh & Schmitt, 2012; Stein, 2010; Taylor, 2008), however, have also noted that morphological awareness of English orthography is needed. Researchers can further pursue this proposal by designing morphologically focused glosses. For instance, target words can be divided based on their root and affixes with textual enhancement of the vowels.

Finally, and as prompted by the shortcomings of the current study, the findings obtained need to be confirmed with further studies by, for example, testing the help options with a higher number of untreated words and with different learner proficiency levels. For instance, advanced learners might show different results with the form-focused glosses than the beginning learners tested in the current study. Ercetin (2013) emphasized that proficiency level influenced the use of help option glosses by ESL learners. In his study, intermediate ESL learners accessed glosses more frequently than the advanced group, however, the advanced group performed better in the reading comprehension task. More generally, the discrepancy in the results of the form-focused glosses for treated and untreated words requires further research to validate the reliability of the current study's findings.

6.6.3 Input enhancement

The input enhancement literature is remarkably consistent in showing nonsignificant findings with respect to the effect of input enhancement in learning target grammatical forms. Nevertheless, the current study, along with the study by Alsadoon and Heift (2015), has provided evidence of the positive effect of input enhancement in the form of highlighting on the reduction of vowel blindness. Both studies suggest that one of the reasons these results differ from the rest of the literature is that it targets a different type of linguistic feature, namely orthographic vowel graphemes. These findings prompt further research to examine the effect of the enhancement on non-grammatical features, given that learning grammar requires intensive and more continuous, long-term practice rather than just one-shot enhancement treatment (Lyddon, 2007).

Input enhancement in the present study was also found successful because the target words were familiar in meaning and form except for learners' confusion with the vowels. Lyddon (2007) emphasized that experimental tasks on input enhancement research in fact trigger global comprehension that distracts learners' attention from the target form. Lee (2007) also indicated that there was a trade-off effect for meaning over form processing in his enhancement study. Accordingly, based on the current study, further research should consider adopting the sequential processing approach as we did, rather than the simultaneous processing approach (Han et al., 2008). Through sequential processing learners' attentional resources are allocated to meaning and then form, hence allowing sufficient time for noticing and processing the target form, especially with input enhancement.

Finally, other enhancement techniques in the context of vowel blindness need to be researched. For instance, other types of textual enhancement can be exploited either explicitly, like typographical enhancement through the choice of different fonts, or implicitly, like input flooding (i.e., embedding many exemplars of the same form in the experimental task), or even a combination of one or more forms of input enhancement. Moreover, the effect of input enhancement on vowel blindness needs to be explored in a less restricted and focused reading context, for instance, in closed captions or video subtitles of a movie.

6.7 Pedagogical Implications

In conclusion, the present study revealed that remedial help in the form of form-focused glosses providing phonics training on segment (grapheme-phoneme) and/or syllable (grapheme-syllable) correspondences, as well as input enhancement, can lead to a significant improvement in reducing vowel blindness decoding errors for Arabic EFL learners. These findings prompt several suggestions for ESL/EFL teachers and the language learning classroom, more generally.

Firstly, teachers must be cognizant of the vowel blindness problem in order to adequately address the problems of Arabic ESL/EFL learners when acquiring their reading, spelling, and writing skills. For this, ESL/EFL schools may offer workshops to

raise the teachers' awareness of the difficulties their learners may encounter when decoding and encoding English vowels.

Secondly, the findings from the study also encourage ESL teachers to raise the awareness of Arabic ESL/EFL learners of their vowel blindness problem and to train them in a natural classroom setting on strategies for word recognition with special focus on vowels. The current study asserts that the majority of study participants were not even aware of their problems with English vowels. Moreover, the study suggests that input enhancement and phonics training can be successful in helping Arabic EFL learners to improve their word recognition skills and acquire missing orthographic vowel knowledge. However, students need to be made aware that such strategies can assist them with their vowel blindness problem. It is also important that teachers implement these strategies from the early stages of ESL/EFL instruction as vowel blindness may delay the learners' progress in acquiring their reading and writing skills.

CALL provides an effective way to implement glossing and input enhancement in the second language classroom. CALL can also deliver other multimedia materials, interactive exercises and feedback through interaction with the computer or with native speakers online. Our study results are based on VALE, which specifically targets vowel blindness. However, highlighting, which turned out to be the most effective help option for both treated and untreated words, can be implemented in any text without computer programming expertise. Similarly, there are a number of authoring tools that readily allow teachers to design and embed glosses similar to those in VALE.

References

Abraham, L. (2007). Second-language reading comprehension and vocabulary learning with multimedia. *Hispania, 90*(1), 98-108.

Abraham, L. (2008). Computer-mediated glosses in second language reading comprehension and vocabulary: A meta-analysis. *Computer Assisted Language Learning, 21*(3), 199-226.

Abu-Rabia, S. (1995). Learning to read in Arabic: Reading, syntactic, orthographic and working memory skills, in normally achieving and poor Arabic readers. *Reading Psychology, 16*(4), 351-394.

Abu-Rabia, S. (2002). Reading in a root-based morphology language: The case of Arabic. *Journal of Research in Reading, 25*(3), 299-309.

Abu-Rabia, S. (2007). The role of morphology and short vowelization in reading Arabic among normal and dyslexic readers in grades 3, 6, 9, and 12. *Journal of Psycholinguistics Research, 36*(2), 89-106.

Abu-Rabia, S., & Awwad, J. (2004). Morphological structures in visual word recognition: The case of Arabic. *Journal of Research in Reading, 25*(3), 299-309.

Abu-Rabia, S., Share, D., & Mansour, M. (2003). Word recognition and basic cognitive processes among reading-disabled and normal readers in Arabic. *Reading and Writing, 16*(5), 423-442.

Abu-Rabia, S., & Taha, H. (2004). Reading and spelling error analysis of native Arabic dyslexic readers. *Reading and Writing, 17*(7-8), 651-690.

Al Ghafli, M. (2011). *The effect of mediated glosses on vocabulary retention and reading comprehension with English language learners in Saudi Arabia* (Unpublished doctoral dissertation). University of Kansas, USA.

Al Mannai, H., & Everatt, J. (2005). Phonological processing skills as predictors of literacy amongst Arabic speaking Bahraini children. *Dyslexia, 11*(4), 269-291.

Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning and Technology, 5*(1), 202-232.

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 259-302). Honolulu, HI: University of Hawaii Press.

Ali, E. M. T. (2011). *Speech intelligibility problems of Sudanese learners of English* (Unpublished doctoral dissertation). Universiteit Leiden, The Netherlands.

Alsadoon, R., & Heift, T. (2015). Textual input enhancement for vowel blindness: A study with Arabic ESL learners. *The Modern Language Journal, 99*(1), 57-79.

Alsulaimani, J. (1990). *Reading problems in Arab learners of English* (Unpublished doctoral dissertation). London University, UK.

Amonette, M. M. (2001). *Beyond the ESL grammar classroom: A descriptive study of transfer of grammatical instruction* (Unpublished doctoral dissertation). University of Washington, USA.

Ariew, R., & Ercetin, G. (2004). Exploring the potential of hypermedia annotations for second language reading. *Computer Assisted Language Learning, 17*(2), 237-259.

Arnbark, E., & Elbro, C. (2002). The effects of morphological awareness training on the reading and spelling skills of young dyslexics. *Scandinavian Journal of Educational Research, 44*(3), 229-251.

Aro, M., & Wimmer, H. (2003). Learning to read: English in comparison to six more regular orthographies. *Applied Psycholinguistics, 24*(4), 621-635.

Awde, N., & Samano, P. (1986). *The Arabic alphabet: How to read and write it*. New York: Kensington Publishing Corp.

Bassetti, B., & Cook, V. (2005). *Second language writing systems*. Clevedon, UK: Multilingual Matters Ltd.

Beland, R., & Mimouni, Z. (2001). Deep dyslexia in the two languages of an Arabic-French bilingual aphasic patients. *Cognition, 82*(2), 77-126.

Blake, R. J. (2009). The use of technology for second language distance learning. *The Modern Language Journal, 93*(1), 822-835.

Bowen, H. (2011). Spelling it out! Accounting for spelling difficulties for Arab learners of English. In T. Smith (Ed.), *Foundations for the future: Focus on vocabulary* (pp. 85-98). United Arab Emirates: HCT Press.

Bowers, P. G., & Wolf, M. (1993). Theoretical links among naming speed, precise timing mechanisms and orthographic skill in dyslexia. *Reading and Writing, 5*(1), 69-85.

Bowles, M. A. (2003). The effects of textual input enhancement on language learning: An online/offline study of fourth-semester Spanish students. In P. Kempchinshky & C. Pineros (Eds.), *In Theory, practice, and acquisition: Papers from the 6th Hispanic linguistic symposium and the 5th conference on the acquisition of Spanish & Portuguese* (pp. 359-411). Summerville, MA: Cascadilla Press.

Bowles, M. A. (2004). L2 glossing: To CALL or not to CALL. *Hispania, 87*(3), 541-552. doi: 10.2307/20063060

Brett, P. (1998). *The design, implementation and evaluation of a multimedia application for second language listening comprehension* (Unpublished doctoral dissertation). University of Wolverhampton, UK.

Brown, T., & Haynes, M. (1985). Literacy background and reading development in a second language. In T. Carr (Ed.), *The development of reading skills* (pp. 19-34). San Francisco: Jossey-Bass.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Language production* (Vol. 2, pp. 257-294). San Diego, CA: Academic Press.

Cardenas-Claros, M. (2005). *Field dependence/field independence: How do students perform in CALL-based listening activities?* (Unpublished masters dissertation). Iowa State University, Ames.

Cardenas-Claros, M. (2011). *A preliminary framework of help options in computer-based second language listening* (Unpublished doctoral dissertation). University of Melbourne, Australia.

Cardenas-Claros, M., & Gruba, P. (2009). Help options in CALL: A systematic review. *CALICO, 27*(1), 69-90.

Chapelle, C. (1996). Validity issues in computer-assisted strategy assessment for language learners. *Applied Language Learning, 7*(1), 47-60.

Chapelle, C. (2001). *Computer applications in second language acquisition*. Cambridge: Cambridge University Press.

Chapelle, C. (2003). *English language learning and technology: Lectures on applied linguistics in the age of information and communication technology*. Amsterdam: John Benjamins.

Chapelle, C. (2005). Hints about CALL use from research. *PacCALL, 1*(1), 1-8.

Chapelle, C. (2009). The relationship between second language acquisition theory and computer-assisted language learning. *The Modern Language Journal, 93*(s-1), 741-753. doi: 10.1111/j.1540-4781.2009.00970.x

Chapelle, C., & Hegelheimer, V. (2000). Methodological issues in research on learner-computer interactions in CALL. *Language Learning and Technology, 4*(1), 41-59.

Chun, D. (2001a). L2 reading on the Web: Strategies for accessing information in hypermedia. *Computer Assisted Language Learning, 14*(5), 367-403.

Chun, D. (2001b). *A longitudinal study of user behavior and L2 reading comprehension in a multimedia CALL environment*. Paper presented at the American Association of Applied Linguistics, St. Louis, MO, February 2001.

Chun, D. (2011). CALL technologies for L2 reading post web 2.0. In N. Arnold & L. Ducate (Eds.), *Present and future promises of CALL: From theory and research to new directions in language teaching* (pp. 131-170). USA: CALICO.

Chun, D., & Payne, J. S. (2004). What makes students click: Working memory and look-up behavior. *System, 32*(4), 481-503.

Chun, D., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *The Modern Language Journal, 80*(2), 183-198.

Chun, D., & Plass, J. L. (1997). Research on text comprehension in multimedia environments. *Language Learning and Technology, 1*(1), 60-81.

Combs, C. (2005). What cognitive processes are triggered by input enhancement? *Teachers College, Columbia University Working Papers in TESOL & Applied Linguistics, 4*(3), 1-14.

Davis, J. (1989). Facilitating effects of marginal glosses on foreign language reading. *The Modern Language Journal, 73*(1), 41-48.

Davis, J., & Lyman-Hager, M. (1997). Computers and L2 reading: Student performance, student attitudes. *Foreign Language Annals, 30*(1), 58-72.

De Ridder, I. (2000). Are we conditioned to follow links? Highlights in CALL materials and their impacts on the reading process. *Computer Assisted Language Learning, 13*(2), 183-195.

De Ridder, I. (2002). Visible or invisible links: Does the highlighting of hyperlinks affect incidental vocabulary learning, text comprehension, and the reading process? *Language Learning and Technology, 6*(1), 123-146.

Doughty, C. (1991). Second language instruction does make a difference. *Studies in Second Language Acquisition, 13*(4), 431-469.

Doughty, C., & Williams, J. (1998). *Focus on form in classroom second language acquisition*. Cambridge, UK: Cambridge University Press.

Ducate, L., & Lomicka, L. (2008). Adventures in the blogosphere: From blog readers to blog writers. *Computer Assisted Language Learning, 21*(1), 9-28.

Ehri, L. (1997). Learning to read and learning to spell are one and the same. In L. Perfetti & M. F. Rieben (Eds.), *Learning to spell: Research, theory, and practice across languages* (pp. 237-268). Mahwah, NJ: Lawrence Erlbaum Associates.

Ercetin, G. (2003). Exploring ESL learners' use of hypermedia reading glosses. *CALICO, 20*(2), 261-283.

Ercetin, G. (2010). Effects of topic interest and prior knowledge on text recall and annotation use in reading a hypermedia text in the L2. *ReCALL, 22*(2), 228-246.

Ercetin, G. (2013). Exploring ESL learners' use of hypermedia reading glosses. *CALICO Journal, 20*(2), 261-283.

Eviatar, Z., & Ibrahim, R. (2000). Bilingual is as bilingual does: Metalinguistic abilities of Arabic-speaking children. *Applied Psycholinguistics, 21*(4), 451-471.

Farahani, A. K., & Sarkhosh, M. (2012). Do different textual enhancement formats have differential effects on the intake of English subjunctive mood? *Theory and Practice in Language Studies, 2*(4), 688-698.

Fender, M. (2003). English word recognition and word integration skills of native Arabic and Japanese speaking learners of English as a second language. *Psycholinguistics, 24*(2), 289-315.

Fender, M. (2008). Spelling knowledge and reading development: Insights from Arab ESL learners. *Reading in a Foreign Language, 20*(1), 19-42.

Figueredo, L. (2006). Using the known to chart the unknown: A review of first-language influence on the development of English-as-a-second language spelling skill. *Reading and Writing, 19*(8), 873-905.

Fox, B., & Routh, D. K. (1975). Analyzing spoken language into words, syllables, and phonemes: A developmental study. *Journal of Psycholinguistic Research, 4*(4), 331-342.

Frost, R. (2005). Orthographic systems and skilled word recognition processes. In M. Snowling & C. Hume (Eds.), *The science of reading: A handbook* (pp. 272-295). Malden, MA: Blackwell.

Frost, R., Forster, K. I., & Deutsch, A. (1997). What can be learned from the morphology of Hebrew? A masked-priming investigation of morphological representation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(4), 829-856.

Fukkink, R. G., Hulstijn, J, H., & Simis, A. (2005). Does training in Second-Language word recognition skills affect reading comprehension? An experimental study. *The Modern Language Journal, 89*(1), 54-75.

Gascoigne, C. (2006). Toward an understanding of incidental input enhancement in computerized L2 environments. *CALICO Journal, 24*(1), 147-162. doi: 10.11139/cj.24.1.147-162

Gass, S., & Mackey, A. (2006). Input, interaction and output: An overview. In K. Bardovi-Harlig & Z. Dörnyei (Eds.), *AILA review* (Vol. 19, pp. 3-17). Amsterdam: John Benjamins Publishing Company.

Gates, L., & Yale, I. (2011). A logical letter−sound system in five phonic generalizations. *The Reading Teacher, 64*(5), 330-339.

Gettys, S., Imhof, L., & Kautz, J. (2001). Computer-assisted reading: The effects of glossing format on comprehension and vocabulary retention. *Foreign Language Annals, 34*(2), 91-106.

Goswami, U., & Bryant, P. (1992). Rhyme, analogy and children's reading. In P. Gough, L. Ehri & R. Treiman (Eds.), *Reading acquisition* (pp. 49-63). Hillsdale, NJ: Lawrence Erlbaum.

Goswami, U., Ziegler, J., Dalton, L., & Scneider, W. (2003). Nonword reading across orthographies: How flexible is the choice of reading units? *Applied Psycholinguistics, 24*(2), 235-247.

Goudarzi, Z., & Moini, M. R. (2012). The effect of input enhancement of collocations in reading on collocation learning and retention of EFL learners. *International Education Studies, 5*(3), 247-258. doi: 10.5539/ies.v5n3p247

Grace, C. (1998). Retention of word meaning inferred from context and sentence-level translations: Implications for the design of beginning-level CALL software. *The Modern Language Journal, 82*(4), 533-544.

Grace, C. (2000). Gender differences: Vocabulary retention and access to translations for beginning language learners in CALL. *The Modern Language Journal, 84*(2), 214-224.

Grgurović, M., & Hegelheimer, V. (2007). Help options and multimedia listening: Student's use of sub-titles and transcripts. *Language Learning and Technology 11*(1), 45-66.

Groff, P. (1971). Dictionary syllabification - How useful? *Elementary School Journal, 72*(3), 900-902.

Groff, P. (1981). Teaching reading by syllables. *The Reading Teacher, 34*(6), 659-664.

Guidi, C. (2009). *Glossing for meaning and glossing for form: A computerized study of the effects of glossing and type of linguistic item on reading comprehension, noticing, and L2 learning* (Unpublished doctoral dissertation). Georgetown University, Washington, DC.

Hamada, M., & Koda, K. (2011). Similarity and difference in learning L2 word-form. *System, 39*(4), 500-509.

Han, Z., Park, E., & Combs, C. (2008). Textual enhancement of input: Issues and possibilities. *Applied Linguistics, 29*(4), 597-618. doi: 10.1093/applin/amn010

Hayes-Harb, R. (2006). Native speakers of Arabic and ESL texts: Evidence for the transfer of the written word identification process. *TESOL Quarterly, 40*(2), 321-339.

Hegelheimer, V. (1998). *Effects of textual glosses and sentence-level audio glosses on online reading comprehension and vocabulary recall* (Unpublished doctoral dissertation). University of Illinois, USA.

Hegelheimer, V., & Tower, D. (2004). Using CALL in the classroom: Analyzing student interactions in an authentic classroom. *System, 32*(2), 185-205.

Heift, T. (2013). Clicking for help. *CALICO Journal, 30*(1), 187-202.

Heift, T., & Chapelle, C. (2011). Language learning through technology. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 555-570). NY: Routledge.

Heikkilä, R., Aro, M., Närhi, V., Westerholm, J., & Ahonen, T. (2013). Does training in syllable recognition improve reading speed? A computer-based trial with poor readers from second and third grade. *Scientific Studies of Reading, 17*(6), 398-414.

Henderson, L., Wallis, J., & Knight, K. (1984). Morphemic structure and lexical access. *Attention and Performance, 10*(1), 211-226.

Hill, M., & Laufer, B. (2003). Type of task, time-on-task and electronic dictionaries in incidental vocabulary acquisition. *International Review of Applied Linguistics, 4*(2), 87-106.

Horst, M. (2005). Learning L2 vocabulary through extensive reading: A measurement study. *The Canadian Modern Language Review, 61*(3), 355-382.

Hubbard, P. (2004). Learner training for effective use of CALL. In S. Fotos & C. Browne (Eds.), *New perspectives in CALL for second language classrooms* (pp. 45-67). Mahwah, NJ: Lawrence Erlbaum.

Hulme, C., Hatcher, J., Nation, K., Brown, A., Adams, J., & Stuart, G. (2002). Phoneme awareness is a better predictor of early reading skill than onset-rime awareness. *Journal of Experimental Child Psychology, 82*(1), 2-28.

Hulstijn, J, H. (1992). Retention of inferred and given word meanings: Experiments in incidental vocabulary learning. In P. Arnaud & H. Bejoint (Eds.), *Vocabulary and applied linguistics* (pp. 113-125). London: Macmillan.

Hulstijn, J, H. (1993). When do foreign-language readers look up the meaning of unfamiliar words? The influence of task and learner variables. *The Modern Language Journal, 77*, 139-147.

Hulstijn, J, H. (2000). The use of computer technology in experimental studies of second language acquisition: A survey of some techniques and some ongoing studies. *Language Learning and Technology, 3*(2), 32-43.

Hulstijn, J, H. (2001). Intentional and incidental second language vocabulary learning: a reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 258-286). Cambridge: Cambridge University Press.

Hulstijn, J, H. (2003). Incidental and intentional learning. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 349-381). Malden, MA: Blackwell.

Hulstijn, J, H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition, 27*(2), 129-140.

Hulstijn, J, H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use, and reoccurrence of unknown words. *The Modern Language Journal, 80*(3), 327-339.

Hulstijn, J, H., & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning, 51*(3), 539-558.

Ibn Manzur, M. (1955). *Lisan al-Arab dictionary*. Beirut: Dar Sadir.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis. *Studies in Second Language Acquisition, 24*(4), 541-577. doi: 10.1017/S0272263102004023

Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition? *TESOL Quarterly, 34*(2), 239-278.

Johnson, A., & Heffernan, N. (2006). The Short Reading Project: A CALL reading activity utilizing vocabulary recycling. *Computer Assisted Language Learning, 19*(1), 63-77.

Johnston, R. S., McGeown, S., & Watson, J. E. (2012). Long-term effects of synthetic versus analytic phonics teaching on the reading and spelling ability of 10 year old boys and girls. *Reading and Writing, 25*(6), 1365-1384.

Jones, L. C., & Plass, J. L. (2002). Supporting listening comprehension and vocabulary acquisition in French with multimedia annotations. *The Modern Language Journal, 86*(4), 546-561.

Jourdenais, R. (1998). *The effects of textual enhancement on the acquisition of the Spanish preterit and imperfect* (Unpublished doctoral dissertation). Ann Arbor, United States. Retrieved from http://search.proquest.com.proxy.lib.sfu.ca/docview/304426364/abstract?accountid=13800

Jourdenais, R., Ota, M., Stauffer, S., Boyson, B., & Doughty, C. (1995). Does textual enhancement promote noticing? A think-aloud protocol analysis. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 183-216). Honolulu: University of Hawaii Press.

Juel, C., Griffith, P., & Gough, P. (1986). Acquisition of literacy: A longitudinal study of children in first and second grade. *Journal of Educational Psychology, 78*(4), 243-255.

Jung, J. (2015). Effects of glosses on learning of L2 grammar and vocabulary. *Language Teaching Research, 19*(4), 473-498.

Karp, A. (2002). *Modification of glosses and its effect on incidental L2 vocabulary learning in Spanish* (Unpublished doctoral dissertation). University of California, Davis.

Katz, L., & Frost, R. (1992). The reading process is different for different orthographies: the Orthographic Depth Hypothesis. In L. Katz & R. Frost (Eds.), *Orthography, phonology, morphology, and meaning* (pp. 67-84). Amsterdam: Elsevier.

Kim, Y. (2003). *Effects of input elaboration and enhancement on second language vocabulary acquisition through reading by Korean learners of English* (Unpublished doctoral dissertation). Ann Arbor, United States. Retrieved from http://search.proquest.com.proxy.lib.sfu.ca/docview/305335411/abstract?accountid=13800

Kitajima, R. (2002). Enhacing higher order interpretation skills for Japanese reading. *CALICO, 19*(3), 571-681.

Koda, K. (2005). *Insights into second language reading: A cross-linguistic approach*. Cambridge: Cambridge University Press.

Koda, K. (2007). Reading and language learning crosslinguistic constraints on second language reading development. *Language Learning, 57*(s-1), 1-44.

Laufer, B., & Hill, M. (2000). What lexical information do L2 learners select in a CALL dictionary and how does it affect word retention? *Language Learning and Technology, 3*(2), 58-76.

Laufer, B., & Hulstijn, J. H. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics, 11*(1), 1-26.

Lee, S. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of the passive form. *Language Learning, 57*(1), 87-118. doi: 10.1111/j.1467-9922.2007.00400.x

Leeman, J. (2003). Recasts and second language development: Beyond negative feedback. *Studies in Second Language Acquisition, 25*(1), 37-63.

Leeman, J., Arteagoitia, I., Fridman, B., & Doughty, C. (1995). Integrating attention to form with meaning: Focus on form in content-based Spanish instruction. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 217-258). Honolulu, HI: University of Hawai'i.

Lemoine, H. E., Levy, B. A., & Hutchinson, A. (1993). Increasing the naming speed of poor readers: Representations formed across repetitions. *Journal of Experimental Child Psychology, 55*(3), 297-328.

Leow, R. (1997). The effects of input enhancement and text length on adult L2 readers' comprehension and intake in second language acquisition. *Applied Language Learning, 8*(2), 151-182.

Leow, R. (1998). Toward operationalizing the process of attention in SLA: Evidence for Tomlin and Villa's (1994) fine-grained analysis of attention. *Applied Psycholinguistics, 19*(1), 133-159.

Leow, R. (2001). Do learners notice enhanced forms while interacting with the L2? An online and offline study of the role of written input enhancement in L2 reading. *Hispania, 84*(3), 496-509.

Leow, R. (2007). Input in the L2 classroom: An attentional perspective on receptive practice. In R. DeKeyser (Ed.), *Practice in a second language: Perspectives from applied linguistics and cognitive psychology* (pp. 21-50). New York, USA: Cambridge University Press.

Leow, R. (2009a). Input enhancement and L2 grammatical development: What the research reveals. In J. Watzinger-Tharp & S. L. Katz (Eds.), *Conceptions of L2 grammar: Theoretical approaches and their application in the L2 classroom* (pp. 16-34). Boston, MA: Heinle Publishers.

Leow, R. (2009b). Input Enhancement in the Classroom. *Hispania, 92*(2), 299-315.

Leow, R., Egi, T., Nuevo, A., & Tsai, Y. (2003). The roles of textual enhancement and types of linguistic item in adult L2 learners' comprehension and intake. *Applied Language Learning, 13*(1), 1-16.

Li, J. (2010). Learning vocabulary via computer-assisted scaffolding for text processing. *Computer Assisted Language Learning, 23*(3), 253-275.

Lin, H., & Chen, T. (2007). Reading authentic EFL text using visualization and advance organizers in a multimedia learning environment. *Language Learning and Technology, 11*(3), 83-106.

Liou, H. (1997). Research of online help as learners strategies for multimedia CALL evaluation. *CALICO Journal, 14*(2-4), 81-96.

Lomicka, L. (1998). "To gloss or not to gloss": An investigation of reading comprehension online. *Language Learning and Technology, 1*(2), 41-50.

Long, M. H. (1996). The role of linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). San Diego, CA: Academic Press.

Lukatela, B., Gligorijevic, A., Kostic, A., & Turvey, M. (1980). Representation of inflected nouns in the internal lexicon. *Memory & Cognition, 8*(5), 415-423.

Lyddon, P. A. (2007). *The efficacy of corrective feedback and target form enhancement in promoting acquisition of the à/au/en/aux distinction in L2 French* (Unpublished doctoral dissertation). University of Arizona, USA.

Lyddon, P. A. (2011). The efficacy of corrective feedback and textual enhancement in promoting the acquisition of grammatical redundancies. *The Modern Language Journal, 95*(s-1), 104-129. doi: 10.1111/j.1540-4781.2011.01272.x

Mahfoudhi, A., Elbeheri, G., Al-Rashidi, M., & Everatt, J. (2010). The role of morphological awareness in reading comprehension among typical and learning disabled native Arabic speakers. *Journal of Learning Disabilities, 43*(6), 500-514.

Marslen-Wilson, W., Tylor, L. K., Waksler, R., & Older, L. (1994). Morphology and meaning in the English mental lexicon. *Psychological Review, 101*(1), 3-33.

Maxim, H. (2002). A study into the feasibility and effects of reading extended authentic discourse in the beginning German language classroom. *The Modern Language Journal, 86*(1), 20-35.

Mayer, E. (1997). Multimedia learning: Are we asking the right question? *Educational Psychologist, 32*(1), 1-19.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (pp. 13–103). New York: Macmillan.

Miles, E. (2000). Dyslexia may show a different face in different languages. *Dyslexia, 6*(3), 193-201.

Milton, J., & Hopkins, N. (2006). Comparing phonological and orthographic vocabulary size: Do vocabulary tests underestimate the knowledge of some learners. *The Canadian Modern Language Review, 63*(1), 127-147.

Morais, J., Cary, L., Alegeria, J., & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously? *Cognition, 7*(4), 323-221.

Muljani, D., Koda, K., & Moates, D. (1998). The development of word recognition in a second language. *Applied Psycholinguistics, 19*(1), 99-113.

Munro, M. J. (1992). *Perception and production of English vowels by native speakers of Arabic* (Unpublished doctoral dissertation), University of Alberta, Canada.

Nadia, J. T., & Charles, W. H. (2011). Contributions of phonological processing skills to reading skills in Arabic speaking children. *Read Write, 24*(9), 1019-1042.

Nagata, N. (1999). The effectiveness of computer-assisted interactive glosses. *Foreign Language Annals, 32*(4), 469-479.

Naser, R. T. (1963). *The teaching of English to Arab students*. London: Longman.

Nation, K., & Hulme, C. (1997). Phonemic segmentation, not onset-rime segmentation, predicts early reading and spelling skills. *Reading Research Quarterly, 32*(2), 154-167.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.

Nikolova, O. R. (2004). Effects of visible and invisible hyperlinks on vocabulary acquisition and reading comprehension for high and average language achievers. *Apprentissage des Langues et Systèmes d'Information et de Communication, 7*(1), 29-53.

Nikolova-Simic, A. (2011). *L1 interference in the perception and production of English vowels by Arabic speakers.* (Unpublished doctoral dissertation), Alliant International University, San Diego.

O'Bryan, A. (2008). Providing pedagogical learner training in CALL: Impact on student use of language-learning strategies and glosses. *CALICO Journal, 26*(1), 142-159.

Odisho, E. Y. (2005). *Techniques of teaching comparative pronunciation in Arabic and English*. Piscataway, NJ: Gorgias Press LLC.

Overstreet, M. (1998). Text enhancement and content familiarity: The focus of learner attention. *Spanish Applied Linguistics, 2*(2), 229-258.

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review 76*(3), 241-263.

Paivio, A. (1986). *Mental representation: A dual-coding approach*. New York: Oxford University Press.

Paradis, M. (1994). Neurolinguistics aspects of implicit and explicit memory: Implications for bilingualism. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 393-419). London: Academic Press.

Park, E., & Nassif, L. (2013). Textual enhancement of two L2 Arabic forms: A classroom-based study. *Language Awareness, 23*(4), 1-19. doi: 10.1080/09658416.2013.808645

Peters, E. (2007). Manipulating L2 learners' online dictionary use and its effects on L2 word retention. *Language Learning and Technology, 11*(2), 36-58.

Pujolà, J. T. (2002). CALLing for help: Researching language learning strategies using help facilities in a web-based multimedia program. *ReCALL, 14*(2), 235-262.

Purnet, J. F., Béland, R., & Idrissi, A. (2000). The mental representation of Semitic languages. *Linguistic Inquiry, 31*(4), 609-648.

Randall, M. (2009). Word recognition, psycholinguistics and teaching second language reading. In D. Anderson (Ed.), *Cultivating real readers* (pp. 13–23). UAE: HCT Press.

Randall, M., & Meara, P. (1988). How Arabs read Roman letters. *Reading in a Foreign Language, 4*(2), 133-145.

Rankin, J. (2005). A case study of embedded extensive reading in intermediate German L2. *Die Unterrichtspraxis, 38*(2), 125-134.

Ravid, D. (2001). Learning to spell in Hebrew: Phonological and morphological factors. *Reading and Writing, 14*(5-6), 459-485.

Robinson, P. (1995). Attention, memory, and the "noticing" hypothesis. *Language Learning, 45*(2), 283-331.

Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition, 18*(1), 27-67.

Roby, W. B. (1999). What's in a gloss? *Language Learning and Technology, 2*(2), 94-101.

Rosa, E., & Leow, R. (2004). Awareness, different learning conditions, and second language development. *Applied Psycholinguistics, 25*(2), 269-292.

Rott, S. (2003). Making form-meaning connections while reading: A qualitative analysis of word processing. *Reading in a Foreign Language, 15*(1), 45-75.

Rott, S. (2005). Processing glosses: A qualitative exploration of how form-meaning connections are established and strengthened. *Reading in a Foreign Language, 17*(2), 95-124.

Rott, S. (2007). The effect of frequency of input-enhancements on word learning and text comprehension. *Language Learning, 57*(2), 165-199. doi: 10.1111/j.1467-9922.2007.00406.x

Ryan, A., & Meara, P. (1991). The case of the invisible vowels: Arabic speakers reading English words. *Reading in a Foreign Language, 7*(2), 531-541.

Sachs, R., & Suh, B-R. (2007). Textually enhanced recasts, learner awareness, and L2 outcomes in synchronous computer-mediated interaction. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: a collection of empirical studies* (pp. 197–227). Oxford, UK: Oxford University Press.

Saiegh-Haddad, E. (2005). Correlates of reading fluency in Arabic: Diglossic and orthographic factors. *Reading and Writing, 18*(6), 559-582.

Saiegh-Haddad, E., & Geva, E. (2008). Morphological awareness, phonological awareness, and reading in English-Arabic bilingual children. *Read Write, 21*(5), 481-504.

Saigh, K., & Schmitt, N. (2012). Difficulties with vocabulary word form: The case of Arabic ESL learners. *System, 40*(1), 24-36.

Saiz, M. (2007). *Does L2 word decoding imply L2 meaning activition? Relationships among decoding, meaning identification, and L2 oral language proficiency in reading Spanish as a second language* (Unpublished doctoral dissertation). University of Pittsburgh, USA.

Sakar, A., & Ercetin, G. (2005). Effectiveness of hypermedia annotations for foreign language reading. *Journal of Computer Assisted Learning, 21*(1), 28-38.

Sanko, G. (2006). The effects of hypertextual input modification on L2 vocabulary acquisition and retention. In M. Nikolov & J. Horvath (Eds.), *Empirical studies in English applied linguistics* (pp. 157-178). Pecs, Hungary: Lingua Franca Csoport.

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics, 11*(2), 129-158.

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics, 13*, 206-226.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *Consciousness in Second Language Learning, 11*(1), 237-326.

Schmidt, R. (Ed.) (1995). *Attention and awareness in foreign language learning*. Honolulu, HI: University of Hawai`i, National Foreign Language Resource Center.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Shalmani, H. B., & Sabet, M. K. (2010). Pictorial, textual, and picto-textual glosses in e-reading: A comparative study. *English Language Teaching, 3*(4), 195-203.

Sharwood Smith, M. (1991). Speaking to many minds: On the relevance of different types of language information for the L2 learner. *Second Language Research, 7*(2), 118-132. doi: 10.1177/026765839100700204

Sharwood Smith, M. (1993). Input enhancement in instructed SLA. *Studies in Second Language Acquisition, 15*(2), 79-165.

Shook, D. J. (1994). FL/L2 reading, grammatical information, and the input-to-intake phenomenon. *Applied Language Learning, 5*(2), 57-93.

Simard, D. (2001). Alertness, orientation, and detection. *Studies in Second Language Acquisition, 23*(1), 103-124.

Simard, D. (2009). Differential effects of textual enhancement formats on intake. *System, 37*(1), 124-135. doi: http://dx.doi.org/10.1016/j.system.2008.06.005

Smith, B. (2001). Arabic speakers. In M. Swan & B. Smith (Eds.), *Learner English: A teacher's guide to interference and other problems* (2nd ed.) (pp. 195-213). Cambridge: Cambridge University Press.

Stanovich, K. E. (2000). *Progress in understanding reading: Scientific foundations and new frontiers*. New York: Guilford Press.

Stein, R. E. (2010). *Sensitivity to consonantal context in reading English vowels: The case of Arabic learners* (Unpublished doctoral dissertation). University of Memphis, USA.

Taft, M. (1981). Prefix stripping revisited. *Journal of Verbal Learning and Verbal Behavior, 20*(3), 289-297.

Taylor, M. (2008). *Orthographic and phonological awareness among L1 Arabic ESL learners: A quasi-experimental study* (Unpublished doctoral dissertation). University of Phoenix, USA.

Thaler, V., Ebner, E. M., Wimmer, H., & Landerl, K. (2004). Training reading fluency in dysfluent readers with high reading accuracy: Word specific effects but low transfer to untrained words. *Annals of Dyslexia, 54*(1), 89-113.

Thompson-Panos, K., & Thomas-Ruzic, M. (1983). The least you should know about Arabic: Implications for the ESL writing instructor. *TESOL Quarterly 17*(4), 609-623.

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition, 16*(2), 183-203.

Tozcu, A., & Coady, J. (2004). Successful learning of frequent vocabulary through CALL also benefits reading comprehension and speed. *Computer Assisted Language Learning, 17*(5), 473-495.

Treiman, R., & Kessler, B. (2005). *Writing systems and spelling development*. Malden, MA: Blackwell.

Truscott, J., & Sharwood Smith, M. (2011). Input, intake, and consciousness: The quest for a theoretical foundation. *Studies in Second Language Acquisition, 33*(4), 497-528. doi: 10.1017/S0272263111000295

Van Gelderen, A., Schoonen, R., Stoel, R., De Glopper, K., & Hulstijn, J, H. (2007). Development of adolescent reading comprehension in language 1 and language 2: A longitudinal analysis of constituent components. *Journal of Educational Psychology, 99*(3), 477-491.

VanPatten, B. (1989). Can learners attend to form and content while processing input? *Hispania, 72*(2), 409-417. doi: 10.2307/343165

VanPatten, B. (1990). Attending to form and content in the input. *Studies in Second Language Acquisition, 12*(3), 287-301. doi: 10.2307/329557

VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Norwood, N.J: Ablex Pub.

VanPatten, B. (2004). *Processing instruction: Theory, research, and commentary*. Nahwah, NJ: Lawrence Erlbaum Associates.

VanPatten, B. (2011). Input processing. In S. Gass & A. Mackey (Eds.), *Handbook of second language acquisition* (pp. 268-281). New York: Routledge.

VanPatten, B., & Cadierno, T. (1993a). Explicit instruction and input processing. *Studies in Second Language Acquisition, 15*(2), 225-243. doi: 10.1017/S0272263100011979

VanPatten, B., & Cadierno, T. (1993b). Input processing and second language acquisition: A role for instruction. *The Modern Language Journal, 77*(1), 45-57. doi: 10.2307/329557

White, J. (1998). Getting the learners' attention: A typographical input enhancement study. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 85-113). New York: Cambridge University Press.

White, L., Spada, N., Lightbown, P. M., & Ranta, L. (1991). Input enhancement and L2 question formation. *Applied Linguistics, 12*(4), 416-432. doi: 10.1093/applin/12.4.416

Winke, P. (2013). The effects of input enhancement on grammar learning and comprehension: A modified replication of Lee (2007) with eye-movement data. *Studies in Second Language Acquisition, 35*(2), 323-352. doi: 10.1017/S0272263112000903

Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. *Language Learning and Technology, 13*(2), 48-67.

Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning and Technology, 10*(3), 85-101.

Yoshii, M, & Flaitz, J. (2002). Second language incidental vocabulary retention: The effect of picture and annotation types. *CALICO, 20*(1), 33-58.

Yun, J. (2011). The effects of hypertext glosses on L2 vocabulary acquisition: A meta-analysis. *Computer Assisted Language Learning, 24*(1), 39-58.

Ziegler, J., Perry, C., Jacobs, A., & Braun, M. (2001). Identical words are read differently in different languages. *Psychological Science, 12*(5), 379-384.

Appendices

Appendix A.

Background Questionnaire

The following questions ask some background information about you. The questionnaire will take a maximum of 15 minutes.

1. Name: …….

2. Age:…….

3. Email: …….

4. Have you studied English abroad? __Yes __No

5. How long have you studied English? _____years _____months

6. How long have you had formal English instruction for reading, speaking, writing, grammar, and listening? _____years _____months

7. When you practice English reading outside the classroom which of the following do you use? Check as many as apply.

__English books

__Internet

__Computer programs

__Social network

__Tutor

__Friend

__Other (please specify)……

8. When you practice English vocabulary outside the classroom which of the following do you use? Check as many as apply.

__English books

__Internet

__Computer programs

__Social networks

__Tutor

__Friend

__Other (please specify)……

9. Do you most often read:

__from the computer?

__print-based materials?

10. What do you prefer:

__to read from the computer?

__to read print-based materials?

11. Do you use the Internet or computer tools for reading such as multimedia, subtitles, reading transcripts, ESL reading comprehension exercises, translation, and other online reading tools?

__Yes __No

If Yes, do you think they are:

__very useful __useful __somewhat useful __ not particularly useful

12. Do you use the Internet or computer tools for learning vocabulary such as dictionaries, word pronunciation, translation, and other online tools?

__Yes __No

If Yes, do you think they are:

__very useful __useful __somewhat useful __ not particularly useful

Thank you!

Appendix B.

Target Words

Target words	Number of letters	Number of syllables
Religion	8	3
Cotton	6	2
Status	6	2
Perfectly	9	3
Weather	7	2
Colours	7	2
Wool	4	1
Friend	6	1
Experience	10	4
Judge	5	1
Assistance	10	3
Decision	8	3
Personal	8	3
Difficult	9	3
Schedule	8	2
Physical	8	3
Muscle	6	2
Relaxed	7	2
Jogging	7	2
Tree	4	1
Practice	8	2
Swimming	8	2
Calm	4	1
Coffee	6	2
Daily	5	2
Step	4	1
Bean	4	1
Roast	5	1
Electric	8	3
Pour	4	1
Sweet	5	1
Winter	6	2
Distractors		
Cough	5	1
Little	6	2

Peace	5	1
Ghost	5	1
Wonderful	9	3
Heat	4	1
World	5	1
Deep	4	1

Appendix C.

Reading Passages

Traditional clothing for men in Saudi Arabia

The religion and customs of Saudi Arabia emphasize modesty in clothing for both men and women. The traditional dress for Saudi men is the “thobe,” which reflects equality regardless of the man’s job or social status. During the summer, Saudi men choose to wear a white thobe which is perfectly suited for the hot weather of Saudi Arabia. During the winter, the thobe is usually a dark colour and made of wool fabric. A man’s headdress consists of three things: a small white cap (taiga), a large square of cloth (the gutra), and a doubled black cord (igal) that holds the gutra in place. The gutra is usually made of cotton and, traditionally, Saudis wear either a white one or a red and white one. However, the white and red gutra is more common.

True friends

True friends will always stand by your side through both the good and the bad. They won't judge you for being you and will give you assistance when you need it. They support you in every decision you make even if they do not agree with you. Most importantly, true friends are always honest with you and tell you the truth no matter what. From my personal experience, it is so difficult to live a life with no friends supporting and encouraging you. I had a hard time during my first four months in Canada with no friends. Now, my true friends make things easier by helping me to enjoy my life and studies. We always have study groups and arrange our study schedules together.

Sports

Sports are good for your physical and mental health. They make the heart muscle strong. Also, they help get rid of toxins from the body. They make you feel relaxed and are good for anxiety. If you feel depressed, go jogging, bike riding or brisk-walking, and you will come back feeling better. This is especially true if you can go somewhere where there are trees and plants to look at so you forget daily difficulties. I think sports are good for your health and brain. I practice sports such as swimming and jogging twice a week and they help me to be calm and stress-free.

Arabian coffee

Arabian coffee is a symbol of strong familial relationships. It has been a daily morning tradition in the Middle East for generations. Usually, the whole family gathers to drink coffee and eat some dates and dried fruits. There are a number of steps in making Arabian coffee. First, prepare fresh green beans to roast either in a coffee roaster or frying pan. Second, grind the coffee beans with a grinder. In the past, they used to use a hand grinder but now almost everyone is using an electric grinder. Third, boil the water to add the coffee grounds. Fourth, leave the coffee to boil for 15 minutes. Fifth, add cardamom and saffron before you take it off the stove. Sixth, pour the hot coffee into a serving coffee pot. Finally, serve it to your guests with dates or any kind of sweets.

Appendix D.

Attitude Questionnaire

Please choose one level of rating for each statement. If you need an Arabic translation help for any of the statements, please do not hesitate to ask your teacher.

Statement	Strongly agree	Agree	Uncertain	Disagree	Strongly disagree
The video tutorial on using VALE provides clear instructions.					
I did not have problems signing up.					
I like the program design, colours, and font size.					
The directory page provides sufficient guidance and instruction.					
The time for each reading session is generally enough.					
I like the optional hide/show feature of the timer.					
The total time of phase 2 (140 minutes) is too much.					
Statement	**Strongly agree**	**Agree**	**Uncertain**	**Disagree**	**Strongly disagree**
Form-focused help option groups					
The information about the vowels is easy to understand.					
The information about the vowels is helpful.					
The information in the glosses seems loaded.					
I had a hard time reading the vowel information in the glosses.					
The vowel information is hard to understand.					
I did not want to check the vowel help options for all the words but VALE forced me to do so.					
I find myself now using the information contained in the vowel help options for other words in English.					
All help option groups					
I did not need any of the vowel help options.					
I noticed that help options improve my word reading, especially for vowels in the words.					
Segment/segment-syllable focused groups					

The audio pronunciation of the word helped me learn the vowel pronunciation of the word.					
The audio pronunciation of vowel sounds helped me read vowel letters in English words.					
Syllable/segment-syllable focused groups					
The syllable types helped me use the vowel information in order to read English words.					
It is easy to read by dividing the word into syllables.					
Dividing the word into syllables helped me remember the word form better.					
Input enhancement group					
I did not understand why some letters are highlighted in yellow.					
I noticed that vowels are highlighted for some words in the reading texts.					
Statement	**Strongly agree**	**Agree**	**Uncertain**	**Disagree**	**Strongly disagree**
I generally did not need to check word meaning.					
I liked the optional access to word meaning.					
I needed to check word meaning for most of the words.					
I needed more information about the word meaning.					
I wish the word meaning were given in English.					
The information contained in the help options helped me realize that I have a problem with reading vowels.					
I do not think that I have a problem with reading vowels.					
I generally confuse vowels and remember consonants.					
I generally feel that I have a problem with remembering word form but I did not know that it was with vowels.					
I have trouble remembering and reading word form but I do think it is due to vowels.					
I noticed that I know the meaning of the words but I could not decide on the correct word form.					

I could not provide the word meaning because I was confused between the word form I chose in the test and the word meaning I know.					

In Arabic or English, tell me what you liked about VALE.

……
……
……
……

In Arabic or English, tell me what you did not like about VALE.

……
……
……
……

Please tell me if you have any other comments or suggestions.

……
……
……
……

Thank you !

Appendix E.

Pre-test

Read each sentence carefully and choose the correct word from in English and word meaning in Arabic from the drop-down menus.

Question	English form	Arabic meaning
1. Snow falls in the …….	wanter, winter, wentir	ابيض, ماء, شتاء
2. Islam is the official …… in Saudi Arabia.	relgion, religon, religion	منطقة, ديانة, توحد
3. My daughter has a terrible…….	cough, caugh, coogh	كحة, كنبة , محكمة
4. Music makes you feel……	relixed, relaxed, rilaxed	مسترخي, مسهل , قوي
5. My school …….. is busy from Saturday to Wednesday.	schedule, sechudle, secuhdle	وحدة تعليمية, جدول , وقت
6. We need a ……house.	little, lettle, lettil	لقب , قليل , قارورة
7. …… have so much sugar.	Sweats, Sweets, Swit	حلويات , تعرق , غرب
8. Language learning requires everyday…….	prictace, practice, prictace	مهارة , اجتهاد , تدريب
9. Red, green, and blue are the basic ……..	coloirs, colaurs, colours	سيارات , ثقافة , الوان
10. The skirt fits me ……...	pirfectly, parfiectly, perfectly	تماماً , غالباً , حالًا
11. The heart is the most important ……..	muscle, muscel, musecle	كتلة , بلح البحر , عضلة
12. Capuccino, Mocca, and Expresso are different kinds of…….	coffe, café, coffee	نعش , قهوة , مقهى
13. I wish they would leave me in …….	piece, peace, pace	قطعة , سلام , سرعة
14. The fridge, washer, and dryer are ……appliances.	elictrecal, electrical, elctrical	كهربائي , ترشيح , اختيار
15. ……. is my favorite sport.	Jugging, Jagging, Jogging	جري , تحكيم , غابة
16. We have a lemon ……..	tree, trea, tria	شاي , شجرة , ربطة
17. Saudi makes Foul from ……..	beens, bains, beans	حبوب , بين , حضر
18. Why do you not ……. a coffee for yourself?	poor, pure, pour	طاهر , فقير , يسكب
19. She needs……. fitness	phyiscial, phiscial, physical	فيزياء , جسدي , فيلسوف
20. …….. stories are scary for me.	Ghost, Gohst, Ghoust	اهداف , شبح , وزة
21. Nama is my best……..	freind, friend, friand	صديق , فريد , مساعد
22. I have …….in the music business.	experience, exprience, expirence	تجربة , توقعات , مجال

23. We had a …….time.	wanderful, wondarful, wonderful	تعجب , رائع , جميل
24. Do you need any……?	assitinance, assistaince, assistance	قوم , تاكيد , مساعدة
25. This is a …… question.	diffecult, defficult, difficult	صعب , يدافع , صلد
26. The …….went off during the snow storm.	hot, heet, heat	حرارة , كرة , قبعة
27. Doctors have high social…….	stutas, status, staitus	تمثال , حالة , ولاية
28. The bank is open ……	dayly, dialy, daily	دليل , يومياً , اتفاق
39. We …….the chicken for dinner.	raost, roast, roost	يحمس , يستريح , يعيد
30. Style is a matter of …….taste.	pirsonal, persanol, personal	شخصي , نفسي , اجتماعي
31. I love the sunny …….	whether, weather, wethear	طقس , اما , بارد
32. I think you have made the right …….	decision, deciosin, decison	وهم , قرار , اتفاق
33. There are a number of …….to make cheesecake.	stips, steps, staps	خطوة , حاد , منحى
34. The weather is …….after a snow storm.	calm, clam, claim	كاميرة , كلام , هادئ
35. I will travel around the …….	werld, wirld, world	كلمة , عالم , ورود
36. Friends do not …….each other.	judge, jadge, jaduge	يحكم , المحلفين , حافة
37. ……. clothes are thick and warm.	Woal, Wale, Wool	عجلة , صوف , جدار
38. I like …….clothes in the summer.	cottan, couttn, cotton	كوت , قطن , قطة
39. ……. is my favorite sport.	Swimming, Swemming, Swomming	سباحة , خياطة , مصارعة
40. Every night, I drink a ……. of water before sleeping.	grass, glass, galss	كاس , عشب , قلص

Appendix F.

Interview Questions

1. What did you like most about VALE?

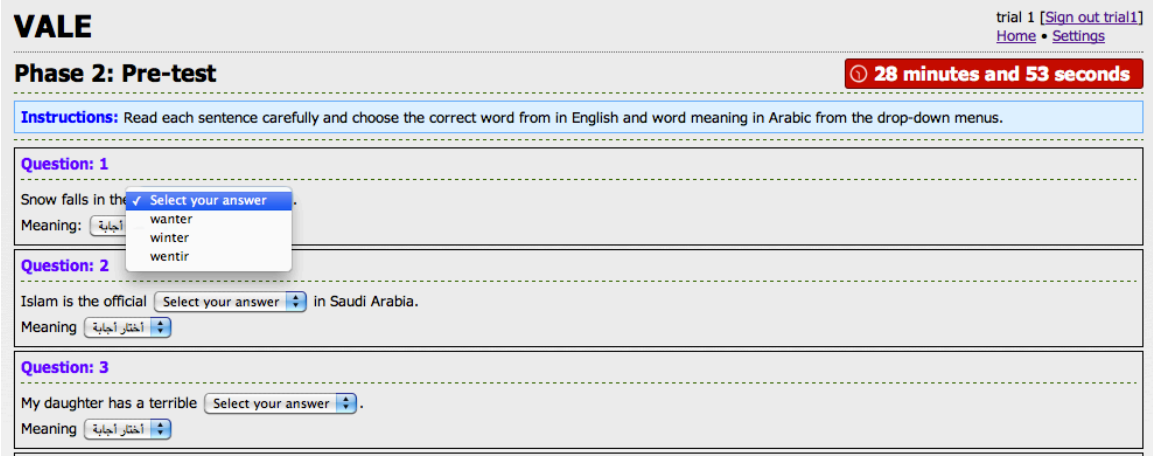
2. What did you like least about VALE?

3. Do you think that VALE helps you improve your reading skills? If so, how?

4. Did you know the words in the tests or were these new words for you?

5. Which part of the test was more difficult for you: choosing the word meaning or the word form? Why?

6. In the tests, you had to choose the correct word form as provided in the picture. What made it difficult for you to choose the correct word?



7. Did you notice that the only difference among the word form choices in the tests concerned vowels?

8. Do you think that the help options in VALE provided enough information to help you notice the vowels?

Form-focused glosses

9. The following picture shows the help options you received during the research study. Which part of the help options was most helpful to you?

Syllable

Religion *(Noun)*

re·li·gion (3 syllables)

Word syllables:	re	li	gion
Syllable types:	Open	Open	Vowel team
Vowel letters/IPA:	ends in e	ends in i	io for the single vowel sound / ə/

Explanation:
An open syllable always ends in a vowel.
A vowel team syllable consists of two vowel letters pronounced as a single sound.

Click for word meaning >>

Segment

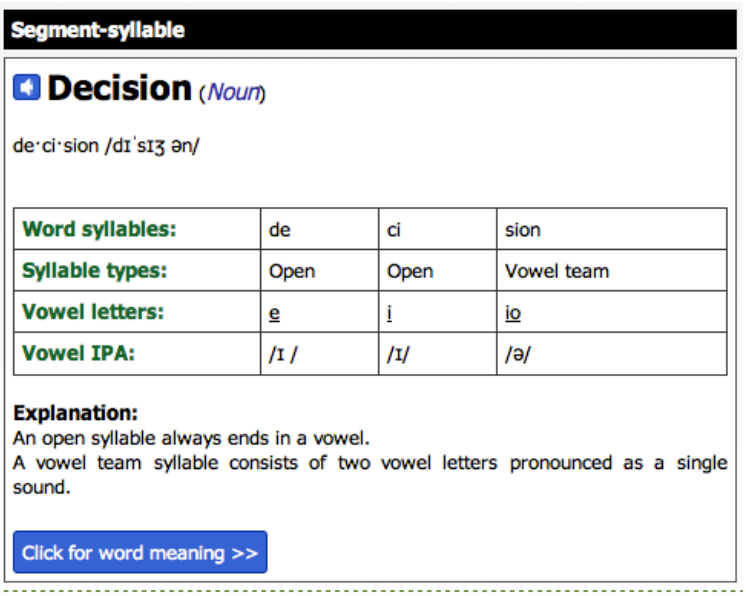
Difficult *(Adjective)*

/dɪf ɪ,kʌlt/

Vowel letters:	i	i	u
Vowel IPA:	/ɪ/	/ɪ/	/ʌ/
Vowel audio:			

Explanation:
The vowel letter i in this word is pronounced as /ɪ/.
The vowel letter u in this word is pronounced as /ʌ/.

Click for word meaning >>



10. Did you notice that the help options focus on vowels?

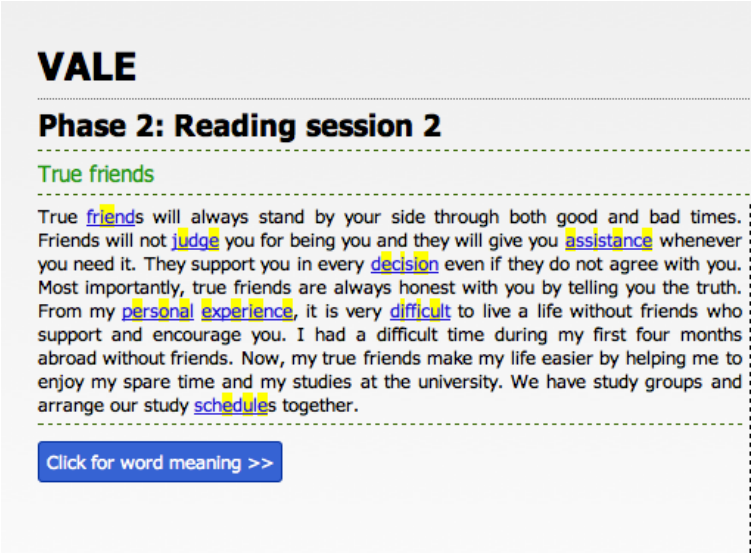
11. What did you like most about the help options? Why?

12. What did you like least about the help options? Why?

13. Do you think that you will be able to apply this knowledge to new words?

Input enhancement

14. The following picture shows an example of the reading passages you received during the study. Do you know what the yellow highlighting stands for?



15. Do you think that the highlighting helped you notice the vowels?

16. Do you think that the highlighting helped you read the word?

17. Do you think that the highlighting provided enough focus on the vowels or would you have preferred additional information?

18. What did you like most about the highlighting? Why?

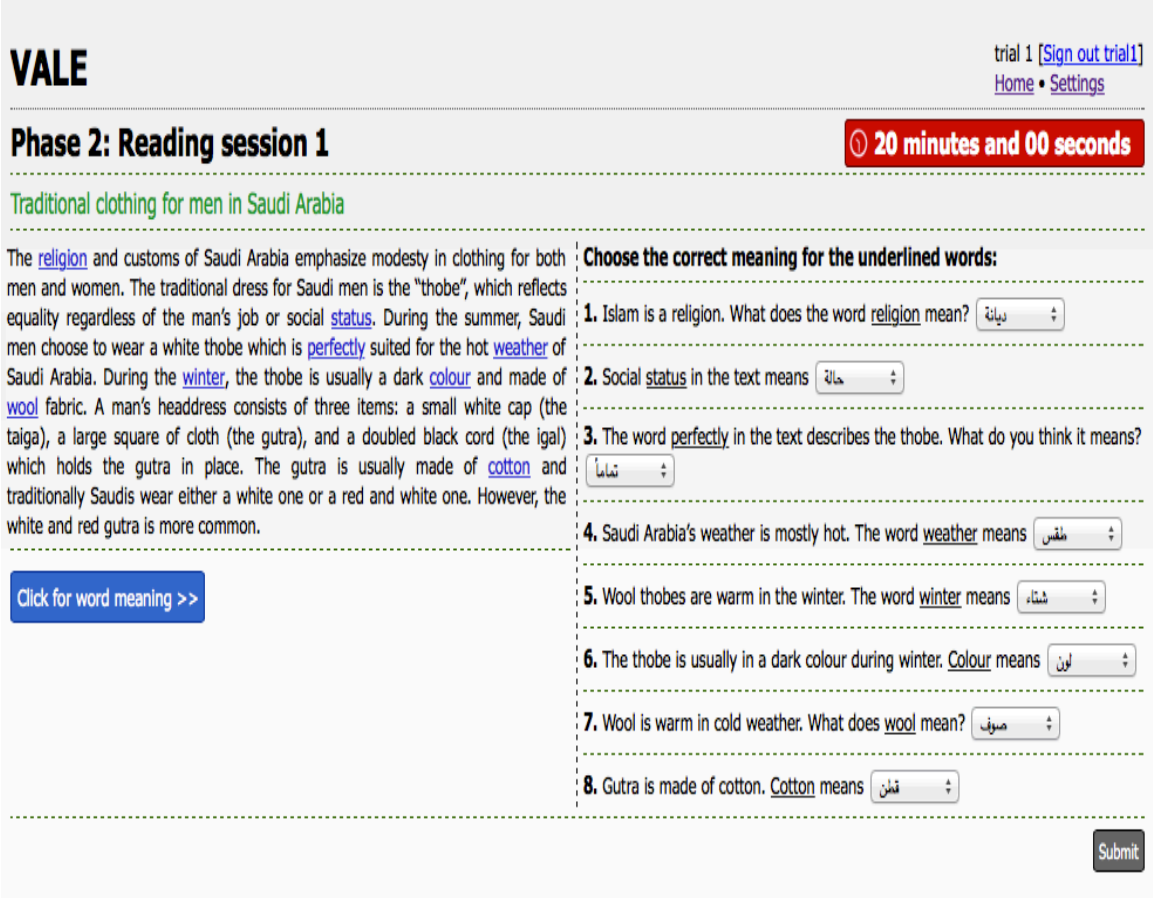
19. What did you like least about the highlighting? Why?

20. The participants of this study were divided into five different groups (*I will show the pictures*). If you had been given the option to choose your group, which one would you have chosen? Why?

21. The reading exercises asked you about the meaning of a word. Were the questions easy for you to answer or were you forced to check the word meanings? Did you check the word meanings often? Why?

Appendix G.

Control condition



Appendix H.

Normality, Homogeneity of Variances, and Kurskall-Wallis Tests on Pre-test scores

Tests of Normality						
Scores	Kolmogorov-Smirnova			Shapiro-Wilk		
	Statistic	Df	Sig.	Statistic	Df	Sig.
Word form	0.100	251	0.000	0.981	251	0.002

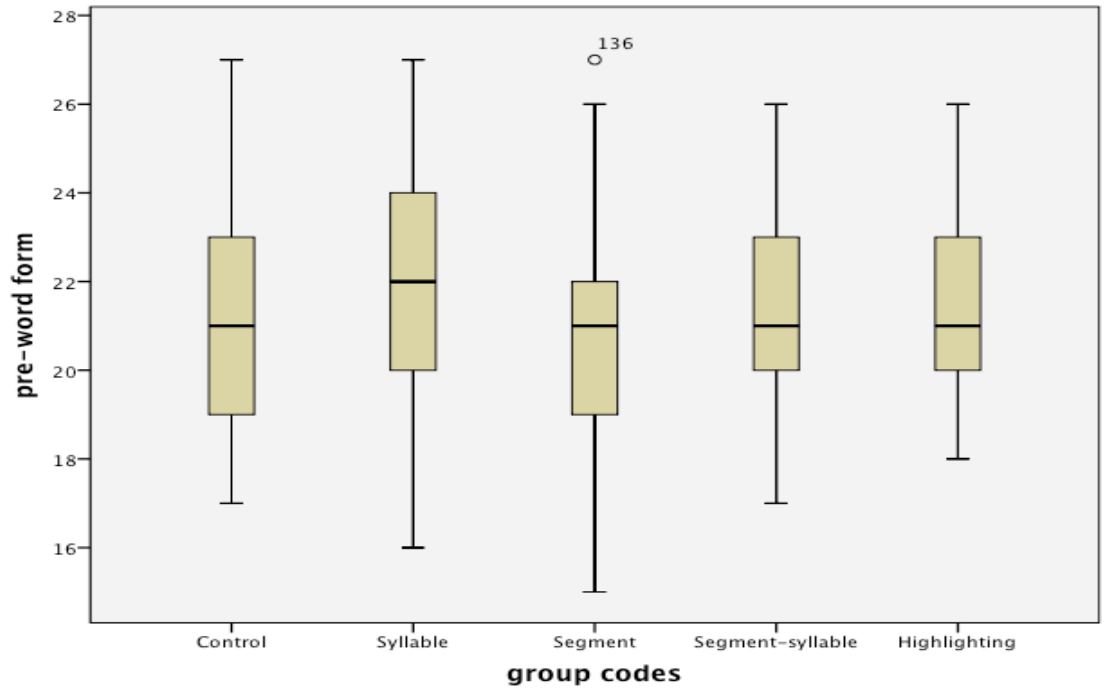
Tests of homogeneity of variances				
Score type	Levene Statistic	df1	df2	Sig.
Word form	1.474	4	246	0.210

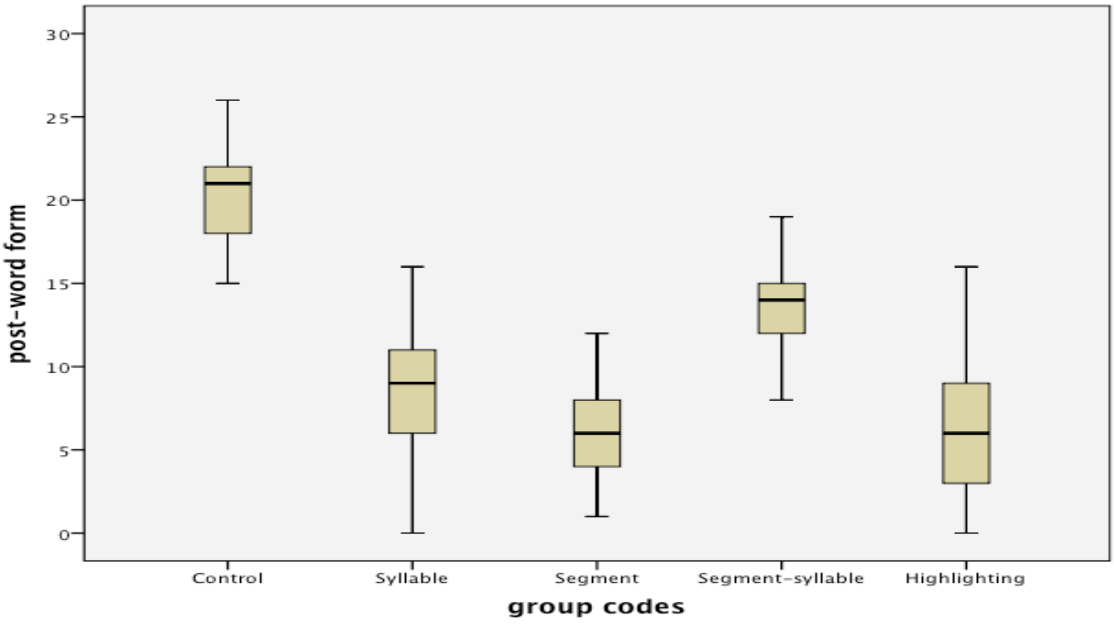
Kruskall-Wallis Tests			
Score type	Chi-Square	Df	Asymp. Sig.
Word form	2.892	4	.576

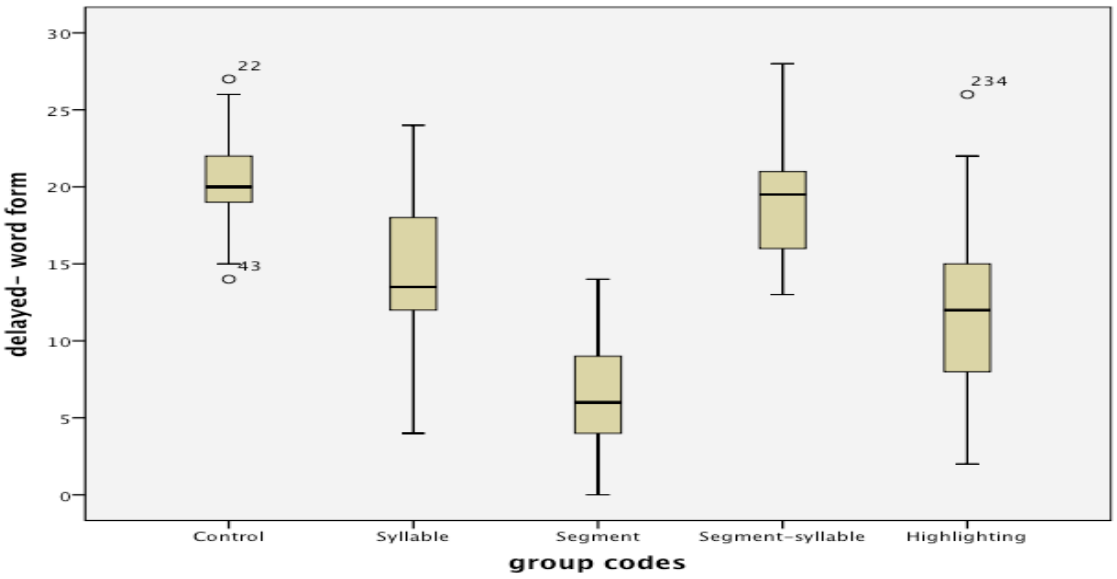
Skewness and Kurtosis analyses on the normality of the data for each group and test

Tests	Groups	Skewness	Kurtosis
Pre-test	Control	.249	-.855
	Syllable	-.192	-.639
	Segment	-.063	.198
	Segment-syllable	-.107	-.472
	Highlighting	.462	-.415
Post-test	Control	.088	-.687
	Syllable	-.285	-.606
	Segment	.133	-.651
	Segment-syllable	-.304	-.114
	Highlighting	.607	-.324
Deylaed post-test	Control	.074	-.439
	Syllable	.047	-.291
	Segment	.392	-755
	Segment-syllable	.181	-.276
	Highlighting	.258	-.006

Plots for Data distribution






delayed- word form
30
25
20
15
10
5
0
22
43
234
Control
Syllable
Segment
Segment-syllable
Highlighting
group codes

Appendix I.

RQ 1.1, 1.2, and 1.3: Tukey Multiple Comparisons of treated words

(I) group codes	(J) group codes	MD (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
			RQ 1.1			
Control	Syllable	12.11843*	0.60285	0.000	10.4617	13.7751
	Segment	13.71843*	0.60285	0.000	12.0617	15.3751
	Segment-syllable	6.77843*	0.60285	0.000	5.1217	8.4351
	Highlighting	13.81843*	0.60285	0.000	12.1617	15.4751
Syllable	Control	-12.11843*	0.60285	0.000	-13.7751	-10.4617
	Segment	1.6	0.60583	0.066	-0.0649	3.2649
	Segment-syllable	-5.34000*	0.60583	0.000	-7.0049	-3.6751
	Highlighting	1.70000*	0.60583	0.043	0.0351	3.3649
Segment	Control	-13.71843*	0.60285	0.000	-15.3751	-12.0617
	Syllable	-1.6	0.60583	0.066	-3.2649	0.0649
	Segment-syllable	-6.94000*	0.60583	0.000	-8.6049	-5.2751
	Highlighting	0.1	0.60583	1.000	-1.5649	1.7649
Segment-syllable	Control	-6.77843*	0.60285	0.000	-8.4351	-5.1217
	Syllable	5.34000*	0.60583	0.000	3.6751	7.0049
	Segment	6.94000*	0.60583	0.000	5.2751	8.6049
	Highlighting	7.04000*	0.60583	0.000	5.3751	8.7049
Highlighting	Control	-13.81843*	0.60285	0.000	-15.4751	-12.1617
	Syllable	-1.70000*	0.60583	0.043	-3.3649	-0.0351
	Segment	-0.1	0.60583	1.000	-1.7649	1.5649
	Segment-syllable	-7.04000*	0.60583	0.000	-8.7049	-5.3751
			RQ 1.2			
Control	Syllable	-5.82314*	0.80762	0.000	-8.0426	-3.6037
	Segment	-0.34314	0.80762	0.993	-2.5626	1.8763
	Segment-syllable	-5.18314*	0.80762	0.000	-7.4026	-2.9637
	Highlighting	-5.10314*	0.80762	0.000	-7.3226	-2.8837
Syllable	Control	5.82314*	0.80762	0.000	3.6037	8.0426
	Segment	5.48000*	0.81161	0.000	3.2496	7.7104
	Segment-syllable	0.64	0.81161	0.934	-1.5904	2.8704
	Highlighting	0.72	0.81161	0.901	-1.5104	2.9504
Segment	Control	0.34314	0.80762	0.993	-1.8763	2.5626
	Syllable	-5.48000*	0.81161	0.000	-7.7104	-3.2496
	Segment-syllable	-4.84000*	0.81161	0.000	-7.0704	-2.6096

	Highlighting	-4.76000*	0.81161	0.000	-6.9904	-2.5296
Segment-syllable	Control	5.18314*	0.80762	0.000	2.9637	7.4026
	Syllable	-0.64	0.81161	0.934	-2.8704	1.5904
	Segment	4.84000*	0.81161	0.000	2.6096	7.0704
	Highlighting	0.08	0.81161	1.000	-2.1504	2.3104
Highlighting	Control	5.10314*	0.80762	0.000	2.8837	7.3226
	Syllable	-0.72	0.81161	0.901	-2.9504	1.5104
	Segment	4.76000*	0.81161	0.000	2.5296	6.9904
	Segment-syllable	-0.08	0.81161	1.000	-2.3104	2.1504
			RQ 1.3			
Control	Syllable	-6.29529*	0.76176	0.000	-8.3887	-4.2019
	Segment	-13.37529*	0.76176	0.000	-15.4687	-11.2819
	Segment-syllable	-1.59529	0.76176	0.226	-3.6887	0.4981
	Highlighting	-8.71529*	0.76176	0.000	-10.8087	-6.6219
Syllable	Control	6.29529*	0.76176	0.000	4.2019	8.3887
	Segment	-7.08000*	0.76552	0.000	-9.1837	-4.9763
	Segment-syllable	4.70000*	0.76552	0.000	2.5963	6.8037
	Highlighting	-2.42000*	0.76552	0.015	-4.5237	-0.3163
Segment	Control	13.37529*	0.76176	0.000	11.2819	15.4687
	Syllable	7.08000*	0.76552	0.000	4.9763	9.1837
	Segment-syllable	11.78000*	0.76552	0.000	9.6763	13.8837
	Highlighting	4.66000*	0.76552	0.000	2.5563	6.7637
Segment-syllable	Control	1.59529	0.76176	0.226	-0.4981	3.6887
	Syllable	-4.70000*	0.76552	0.000	-6.8037	-2.5963
	Segment	-11.78000*	0.76552	0.000	-13.8837	-9.6763
	Highlighting	-7.12000*	0.76552	0.000	-9.2237	-5.0163
Highlighting	Control	8.71529*	0.76176	0.000	6.6219	10.8087
	Syllable	2.42000*	0.76552	0.015	0.3163	4.5237
	Segment	-4.66000*	0.76552	0.000	-6.7637	-2.5563
	Segment-syllable	7.12000*	0.76552	0.000	5.0163	9.2237

Appendix J.

RQ 2.1, 2.2, and 2.3: Tukey Multiple Comparisons of untreated words

(I) group codes	(J) group codes	MD (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
			RQ 2.1			
Control	Syllable	1.48510*	0.22641	0.000	0.8629	2.1073
	Segment	1.22510*	0.22641	0.000	0.6029	1.8473
	Segment-syllable	1.48510*	0.22641	0.000	0.8629	2.1073
	Highlighting	1.72510*	0.22641	0.000	1.1029	2.3473
Syllable	Control	-1.48510*	0.22641	0.000	-2.1073	-0.8629
	Segment	-0.2600	0.22753	0.784	-0.8853	0.3653
	Segment-syllable	0.00000	0.22753	1.000	-0.6253	0.6253
	Highlighting	0.2400	0.22753	0.829	-0.3853	0.8653
Segment	Control	-1.22510*	0.22641	0.000	-1.8473	-0.6029
	Syllable	0.2600	0.22753	0.784	-0.3653	0.8853
	Segment-syllable	0.2600	0.22753	0.784	-0.3653	0.8853
	Highlighting	0.5000	0.22753	0.184	-0.1253	1.1253
Segment-syllable	Control	-1.48510*	0.22641	0.000	-2.1073	-0.8629
	Syllable	0.0000	0.22753	1.000	-0.6253	0.6253
	Segment	-0.2600	0.22753	0.784	-0.8853	0.3653
	Highlighting	0.2400	0.22753	0.829	-0.3853	0.8653
Highlighting	Control	-1.72510*	0.22641	0.000	-2.3473	-1.1029
	Syllable	-0.2400	0.22753	0.829	-0.8653	0.3853
	Segment	-0.5000	0.22753	0.184	-1.1253	0.1253
	Segment-syllable	-0.2400	0.22753	0.829	-0.8653	0.3853
			RQ 2.2			
Control	Syllable	-0.49961	0.23968	0.230	-1.1583	0.159
	Segment	-0.59961	0.23968	0.094	-1.2583	0.059
	Segment-syllable	-0.27961	0.23968	0.770	-0.9383	0.379
	Highlighting	0.34039	0.23968	0.615	-0.3183	0.999
Syllable	Control	0.49961	0.23968	0.230	-0.1590	1.1583
	Segment	-0.1000	0.24086	0.994	-0.7619	0.5619
	Segment-syllable	0.2200	0.24086	0.892	-0.4419	0.8819
	Highlighting	.84000*	0.24086	0.005	0.1781	1.5019
Segment	Control	0.59961	0.23968	0.094	-0.0590	1.2583
	Syllable	0.1000	0.24086	0.994	-0.5619	0.7619
	Segment-syllable	0.3200	0.24086	0.674	-0.3419	0.9819

	Highlighting	.94000*	0.24086	0.001	0.2781	1.6019
Segment-syllable	Control	0.27961	0.23968	0.770	-0.3790	0.9383
	Syllable	-0.2200	0.24086	0.892	-0.8819	0.4419
	Segment	-0.3200	0.24086	0.674	-0.9819	0.3419
	Highlighting	0.6200	0.24086	0.078	-0.0419	1.2819
Highlighting	Control	-0.34039	0.23968	0.615	-0.9990	0.3183
	Syllable	-.84000*	0.24086	0.005	-1.5019	-0.1781
	Segment	-.94000*	0.24086	0.001	-1.6019	-0.2781
	Segment-syllable	-0.6200	0.24086	0.078	-1.2819	0.0419
			RQ 2.3			
Control	Syllable	.98549*	0.18561	0.000	0.4754	1.4956
	Segment	.62549*	0.18561	0.008	0.1154	1.1356
	Segment-syllable	1.20549*	0.18561	0.000	0.6954	1.7156
	Highlighting	2.06549*	0.18561	0.000	1.5554	2.5756
Syllable	Control	-.98549*	0.18561	0.000	-1.4956	-0.4754
	Segment	-0.3600	0.18653	0.304	-0.8726	0.1526
	Segment-syllable	0.2200	0.18653	0.763	-0.2926	0.7326
	Highlighting	1.08000*	0.18653	0.000	0.5674	1.5926
Segment	Control	-.62549*	0.18561	0.008	-1.1356	-0.1154
	Syllable	0.3600	0.18653	0.304	-0.1526	0.8726
	Segment-syllable	.58000*	0.18653	0.018	0.0674	1.0926
	Highlighting	1.44000*	0.18653	0.000	0.9274	1.9526
Segment-syllable	Control	-1.20549*	0.18561	0.000	-1.7156	-0.6954
	Syllable	-0.2200	0.18653	0.763	-0.7326	0.2926
	Segment	-.58000*	0.18653	0.018	-1.0926	-0.0674
	Highlighting	.86000*	0.18653	0.000	0.3474	1.3726
Highlighting	Control	-2.06549*	0.18561	0.000	-2.5756	-1.5554
	Syllable	-1.08000*	0.18653	0.000	-1.5926	-0.5674
	Segment	-1.44000*	0.18653	0.000	-1.9526	-0.9274
	Segment-syllable	-.86000*	0.18653	0.000	-1.3726	-0.3474