Lexical access by different script adult bilinguals; evidence from masked primed picture naming and phoneme monitoring tasks.

Manal Alharbi

Thesis submitted in partial fulfilment of the degree of Doctor of Philosophy

Centre for Language and Communication Research
School of English, Communication and Philosophy Cardiff University

May 2022

Abstract

This study investigates lexical access in different script bilinguals; namely Arabic-English adult speakers, a group that is rarely investigated. It covers cross-language activation, the manner of lexical/phonological selection, and the flow of activation in different script bilinguals, and whether proficiency level modulates cross-language activation, manner of lexical/phonological selection, or flow of activation.

Currently there is a substantial and rapidly growing body of empirical evidence describing and evaluating lexical access in bilinguals. However, the majority of these studies focus on same script bilinguals (e.g., Spanish-English), with limited research addressing different script bilinguals, such as Arabic-English bilinguals. It is argued here that the findings concerning non-selective lexical access cannot be generalized to different script bilinguals, as script differences can act as language cues and strict lexical access to target language. Therefore, we conducted five different experiments to investigate the performance of highly and less proficient adult Arabic-English bilinguals using three different tasks: a masked primed picture naming task (experiments one, two, and three), a phoneme monitoring task (experiments four and five), and an animacy decision task (experiment three). The use of the masked priming paradigm to address lexical access, manner of selection and flow of activation in bilinguals is unprecedented.

In the first experiment, the participants were required to name in the L2 (English), cognate and non-cognate pictures that were preceded by L1 (Arabic) masked translation primes. A significant cognate facilitation effect, and a translation facilitation effect were observed for both highly and less proficient bilinguals. These findings suggest the bilinguals experienced non-selective access (i.e., both languages were activated simultaneously), and that the manner

of lexical/phonological selection was language specific (i.e., considers activated nodes in the target language only). Moreover, the findings pertaining to cognate facilitation suggest a cascaded flow of activation for the non-target language. In experiment two, non-cognate pictures were preceded by L1 semantically related masked primes, which were presented for 50 ms, 75 ms, and 100 ms. The semantic interference effect was evident when the masked primes were presented for 75 ms, and 100 ms. This suggests that the lexical selection process is language non-specific (i.e., considers activated nodes in the target and non-target language), which contradicts the findings reported in experiment one. Experiment three investigated the locus of the semantic interference effect to establish if it is at the conceptual or lexical level. Thus, it compared the effect of semantically related masked primes across two tasks; i.e., the animacy decision task, which involves conceptual processing, and the masked primed picture naming task, which involves both conceptual and lexical processing. The results demonstrated that the semantic interference effect was obtained in the masked primed picture naming task; whereas semantic facilitation effect was obtained in the animacy decision task. This suggests the locus of semantic effect is at the lexical level. In experiment four, the participants performed a phoneme monitoring task in L2, in which they had to decide whether a visually presented phoneme was part of the L2 picture name. The phonemes were either, part of the picture name in L2 in the positive condition, part of the picture name in L1 in the critical condition, or unrelated. The results revealed the participants experienced difficulties rejecting the phoneme when it was part of the L1 picture name. Thus, the Arabic-English bilinguals' two languages appear to be activated simultaneously, and the activation of the non-target language cascades to the phonological level. Experiment five examined whether the findings of the phoneme monitoring task in experiment four would be replicated when L1 distinct phonemes (do not exist in participants' L2 language) were used in an additional critical condition. It was found to be hard to reject L1 phonemes, even when the phonemes are L1 distinct. The findings confirm non-selective access, as well as the cascaded flow of activation during the production process by different script bilinguals.

Taken together, the results reported suggest the manner of lexical access in different script bilinguals is language non-specific, and that activation flow cascades to the phonological level. In addition, the results imply the lexical/phonological selection process considers the activation of the target and non-target lexical nodes. Regarding the role of script differences and the participants' language proficiency level, the results suggest no modulation of cross-language activation, manner of lexical/phonological selection, or flow of activation. The implications for bilingual models of lexical access are also discussed.

Acknowledgement

I would like to express my gratitude to all those who gave me the possibility to complete this thesis. I am deeply indebted to my supervisor Dr. Michelle Aldridge-Waddon whose help, stimulating suggestions and encouragement helped me in all the time of research and writing of this thesis. Special thanks to my parents whose efforts brought me to where I am today.

This dissertation is dedicated to my husband Waleed Ghonaim whose love, endless support and encouragement enabled me to finish this work. I am heartily thankful to every member in my family for being a great help in difficult times. To my two little angels Khalid and Sarah, who will hopefully read this when they grow up and understand what kept their Mum busy during their early childhood years.

To Albert Costa (1970-2018), the talented researcher whose seminal work addressed key questions in bilingualism. His work has inspired me and many psycholinguistic researchers and will continue inspiring future research students. His reader-friendly reviews attracted me to the field of psycholinguistics.

Lastly, I offer my regards and blessings to all of those who supported me in any respect during the completion of the project.

TABLE OF CONTENT

List of Figures

List of Tables

Chapter 1 Introduction and Background

The main goal of this research is to investigate lexical access by different script (Arabic-English) adult bilinguals and the manner of lexical/phonological selection, namely whether it is language specific or non-specific. It also seeks to explore the flow of activation, that is whether the activation of the non-target language cascades to the phonological level. In brief, the research focuses on how adult bilinguals select the target word when lexical items from both languages are likely to be available. The second goal of the research is to test whether participants' language proficiency level modulates lexical access, the manner of lexical/phonological selection, and the flow of activation in different script bilinguals. The motivation for this research is that it offers valuable insights into our theoretical understanding of lexical production models (cf. chapters four, five, six, seven, eight, and nine) and has interesting practical applications (cf. chapters four, five, six, seven, and eight). Moreover, by researching the lexical access process by two different proficiency groups (highly and less proficient) in their second language (L2) across five experiments, it is possible to critically evaluate how processing changes with language exposure and practice, and thereby the study's findings may have interesting implications for pedagogy.

Word production is defined as the process by which we translate concepts and thoughts into patterns of sounds using our articulatory organs (Costa, Colomé, et al. 2000). Abstract nodes are stored in our mental lexicon, and lexical access describes the process of selecting items for production. For most adult monolingual speakers there are multiple items available to label an object (e.g., toy, teddy, comforter), and of course for bilingual speakers there are potentially twice as many items, leading to the obvious question of how the target item is accessed in real time, and what factors inhibit or enhance the process of making a connection between the concept and the corresponding target lexical node (La Heij 2005).

In the published literature (e.g., Costa and Caramazza 1999; Hermans et al. 1998), there is compelling evidence that the activated concept (i.e., the intended message; for example, the concept of '*a dog'*) sends activation to the corresponding lexical nodes in the target and non-target language (namely the two alternatives '*dog'* and '*perro'* for an English-Spanish bilingual). The selection process is achieved through a lexical selection mechanism that selects the target lexical node in the target language. However, whether or not the non-target item is also considered for selection is heavily debated. Currently, there are two opposing views in the literature regarding the manner of lexical selection, namely language specific selection and language non-specific selection. The language specific selection view (Roelofs et al. 1998; Costa and Caramazza 1999; Costa et al. 1999) postulates that only the lexical nodes in the target language are considered during lexical selection, whereas the language non-specific selection view (De Bot 1992; Green 1998; Hermans et al. 1998) postulates that lexical nodes in the target and non-target language are considered during selection. Another point of disagreement in the literature concerns the flow of activation of the non-target lexical nodes, namely whether it cascades from the lexical level to the phonological level (cascaded view) (Caramazza 1997; Costa, Caramazza, et al. 2000; Gollan and Acenas 2000), or whether it does not go beyond the lexical level, and thus does not cascade to the phonological level (the discrete view) (Levelt 1989). Most recent findings of same script bilingual studies suggested a cascaded view, however whether the same cascaded flow is applicable to different script bilinguals remains unclear (see for review Costa 2004). The objective of the present study is therefore to determine whether testing adult bilinguals with different scripts can shed light on these debates, and reveal the underlying processing of lexical access, manner of selection, and activation flow further.

The underlying mechanism of lexical access in bilinguals is typically described as complex and complicated as it involves pragmatic, semantic, syntactic, phonological, and articulatory processes and presentations (Costa, Santesteban, et al. 2006). The present study illustrates how the complex nature of bilingual lexical access and the reported methodological limitations of tasks, such as the often-used picture word interference task (Miozzo and Caramazza 2003; Costa, La Heij, et al. 2006) have contributed to the contradictory findings in the literature regarding the manner of lexical/phonological selection and the flow of activation. This study identifies a need to shift to different behavioural paradigms, in order to generate more informative data to unravel the current contradictions and develop further known theories. Thus, this research deliberately deviates from typical studies concerning bilingual lexical access reported in the literature, as it introduces a different experimental procedure that adopts the masked primed picture naming and the animacy decision with the phoneme monitoring tasks, and works with adults using different scripts, rather than the more traditional same script bilinguals.

There is currently a substantial body of empirical data describing lexical access by bilinguals and monolinguals (Green 1986; Levelt 1989; Schriefers et al. 1990; Roelofs 1992; Hermans et al. 1998; Costa and Caramazza 1999). However, the vast majority of these studies examined same script bilinguals (e.g., Spanish-English, Spanish-Catalan, and French-English bilinguals), and little attention was paid to different script bilinguals, such as Arabic-English bilinguals. As discussed later in chapter two, a shared script may contribute to the findings of non-selective access, and there is a possibility that a different pattern of access may manifest if the languages have a distinct script. Previous research on same script bilinguals (e.g., Miller and Kroll 2002) reported that bilinguals use language specific features as a language cue that in turn reduces the cross-language activation during production. Script differences are prominent linguistic features, thus we expect that these differences may act as a language cue and modulate lexical access/activation and the selection process, because, as detailed later in chapter two, it is known that orthography plays an important role in word production, even when the written form is absent (e.g., Han and Choi 2016). Thus, this study seeks to further the understanding of bilingual lexical access by investigating different script bilinguals (Arabic-English speakers) whose two languages differ substantially, especially in the written form. This research is not the first to examine different script bilinguals, and the work of scholars such as Hoshino (2006), Moon and Jiang (2012), and Kheder and Kaan (2019) is reviewed later in this thesis. However, these studies adopted identical experimental paradigms to those used when testing same-script bilinguals, namely simple picture naming, picture-word interference tasks, and language switching, while the present study argues that these paradigms are unable to adjudicate between the language-specific and language non-specific hypothesis (Costa, Colomé, et al. 2000), as detailed in chapter two, section 2.2.2. Therefore, the current study employs a different paradigm, namely the masked priming paradigm, as the main task in experiments one, two and three (chapters four, five and six, respectively). In addition, Arabic-English bilinguals are recruited to this study, as the population has rarely participated in bilingual lexical access studies. Moreover, only a few previous studies examined the effect of L2 proficiency level on lexical access in bilingual speakers (e.g., Boukadi et al. 2015), with mixed results. The findings of the current study will therefore fill this gap, and clarify some of the previous contradictory findings by analysing the data from two distinct proficiency level groups (highly and less proficient) in five different experiments (chapters four, five, six, seven, and eight).

1.1 General objectives of the study

In brief the general objectives of this research are :

- To determine whether or not lexical access in different script bilinguals is language non-specific;
- To investigate the manner of lexical/phonological selection and the flow of activation during bilingual word production;
- To examine whether cross-language activation, manner of lexical/phonological selection, and flow of activation is influenced by proficiency level.

The study aims to:

- Further our understanding of adult lexical access, in order to develop the current word production models;
- Evaluate the role of orthography and proficiency level in this process;
- Conclude whether it is possible to make any recommendations for second language learning pedagogy.

1.2 Research questions

The current research attempts to answer the following questions:

- What is the manner of lexical access in different script bilinguals?
- What is the manner of lexical selection in different script bilinguals?
- Does the flow of activation cascade to the phonological level in different script bilinguals?
- What is the manner of phonological selection in different script bilinguals?
- What is the effect of proficiency level on the manner of lexical access, lexical/phonological selection, and flow of activation in different script bilinguals?

1.3 Organizations of the thesis

This thesis is organized in nine chapters. The first chapter has introduced the topic and the objectives of the study. The second chapter presents a literature review of lexical access in bilinguals, including the proposed models of word production, and a review of the traditional methods employed in such investigations. It also introduces the masked priming paradigm and the animacy decision task that are employed by the present study, and justifies their adoption. The third chapter summarizes the five different experiments conducted in this research, and the methods implemented in each task. The next chapters (four, five, six, seven, and eight) describe each experiment in turn, in detail, reporting and discussing the results and the implications of their findings. Finally, the last chapter provides a general discussion of the overall findings and their implications for bilingual models of language production. It also briefly discusses the pedagogical implications of the study's findings, and the limitations of the present research.

Chapter 2 Literature Review

This chapter commences with a concise overview of the preliminary theories of lexical access by monolinguals, regarding the manner of activation and selection (section 2.1), in order to facilitate readers' understanding of the processes involved, before shifting to the discussion of lexical access by bilingual speakers discussed in the second section (2.2). The discussion expands in section 2.3 to include a critical evaluation of the common methods used in such investigations. Section 2.3.5 then highlights the fact that lexical access in different script bilinguals has received little attention to date, as well as the role of proficiency level in bilinguals' word production (discussed in section 2.4). This is followed by a discussion of the implications of this gap in the extant literature. The next section (2.5) presents a brief review of the role of orthography in spoken word production, despite the absence of the written form. Section 2.6 discusses the role of language cues, and how specific linguistic features of any language can influence the production process of similar/different script bilinguals. An overview of the cross-linguistic differences between the Arabic and English language is introduced in section 2.7. The chapter concludes with a detailed account of the priming method (section 2.8), highlighting the ways in which this paradigm is suitable for investigating word production by bilinguals. This is followed in section 2.9 by a brief review of the animacy decision task, detailing how it helps to address the research questions of the present study.

2.1 An overview of lexical access in speech production in monolinguals: Stages and processes

Early word production studies (Levelt 1989; Levelt et al. 1999) investigated the architecture and process of the production system by monolingual speakers, and initially theoretical models were developed (e.g., Dell 1986; Levelt 1989; Levelt et al. 1999) that captured the production process of monolingual speakers. Subsequent studies (e.g., Green 1986; De Bot 1992; Poulisse and Bongaerts 1994; Poulisse 1997) investigated whether these models could be extended to bilingual speakers, which were the main focus of the present study. Since most bilingual models are based on proposals originally made in relation to monolingual lexical access, this section introduces these models and their predictions before proceeding to review the existing bilingual lexical access models. The main questions addressed in the previous psycholinguistics research concerning monolingual word production (Levelt 1989; Levelt et al. 1999) were as follows:

(i) How many stages of processing (i.e., levels of representation) are there in speech production?

(ii) Are these levels independent of one another, or is there an interaction between them?

(iii) Does only the target[1] word receive activation, or do other related, non-target words also receive activation, and then compete for selection?

(iv) If competition occurs, how do speakers select the target word?

Previous studies proposed different models of speech production to address these questions (cf. Dell 1986; Levelt 1989; Caramazza 1997; Dell et al. 1997; Levelt et al. 1999). These models recognized at least three stages of processing: the conceptual level, where the meanings of the words are stored; the lexical/lemma level, where the syntactic/semantic properties of the words are stored; and the lexeme, in which information about the word forms (morpho-phonological properties) is stored (Levelt 1989; Levelt et al. 1991; Roelofs 1992; Bock and Levelt 1994) (Figure 1).

[1] Throughout the study, I make use of the following terms 'target word' and 'target lexical nodes' interchangeably. In monolingual studies, these terms refer to the specific word that the speaker wishes to produce to name an object. In bilingual studies, they refer to the specific word in the language in which the speaker is performing the task.

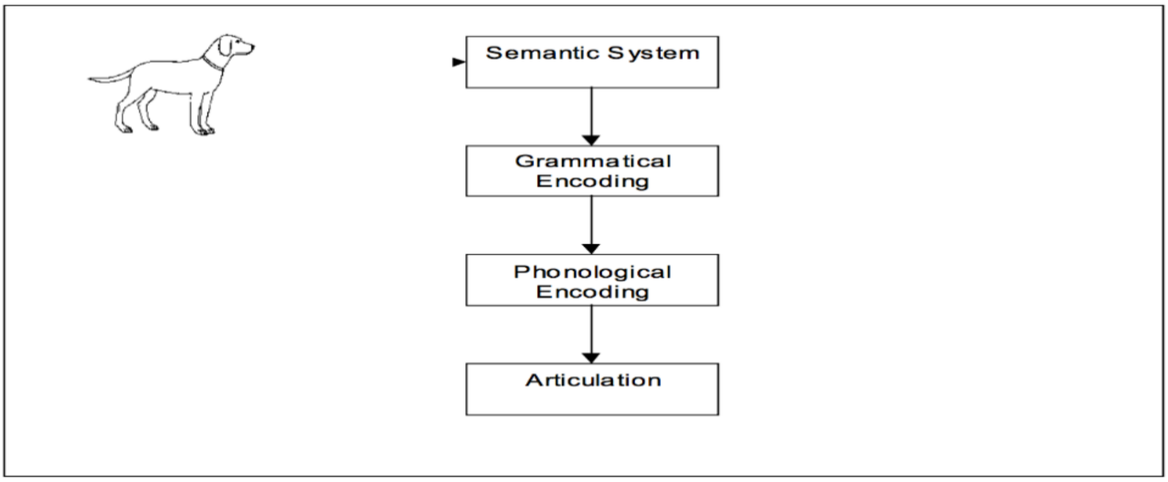


Figure. 1 The different stages of processing involved in speech production (Costa, Colomé et al. 2000, p. 406).

To illustrate, when a speaker attempts to name an object, such as *a dog*, it is argued that not only the concept of '*a dog*'[2] is activated, but several semantic representations (such as '*cat*') also receive activation to some level, either because they share some semantic features, such as being a four-legged animal (Levelt 1989), or because their semantic representations are interconnected (Dell 1986; Caramazza 1997) (see Figure 2). These activated semantic representations spread activation to the corresponding lexical nodes at the lexical/grammatical encoding level, which is often referred to as 'the spreading activation principle'[3] (Dell 1986; Starreveld and La Heij 1995; Caramazza 1997; Roelofs et al. 1998; Levelt et al. 1999; Costa, Colomé, et al. 2000). The speaker then must select the target lexical node corresponding to the picture of *a dog* among the activated non-target lexical nodes (such as '*a cat*').

---

[2] Throughout this study, the following notation is used: italics for stimuli (pictures or words), italics and single quotation marks for lexical and semantic representations, and round brackets for the meaning of the stimuli (picture or words) if it is presented in any language other than English.

[3] The spreading activation principle refers to the activation that spreads from conceptual system to the lexical level.

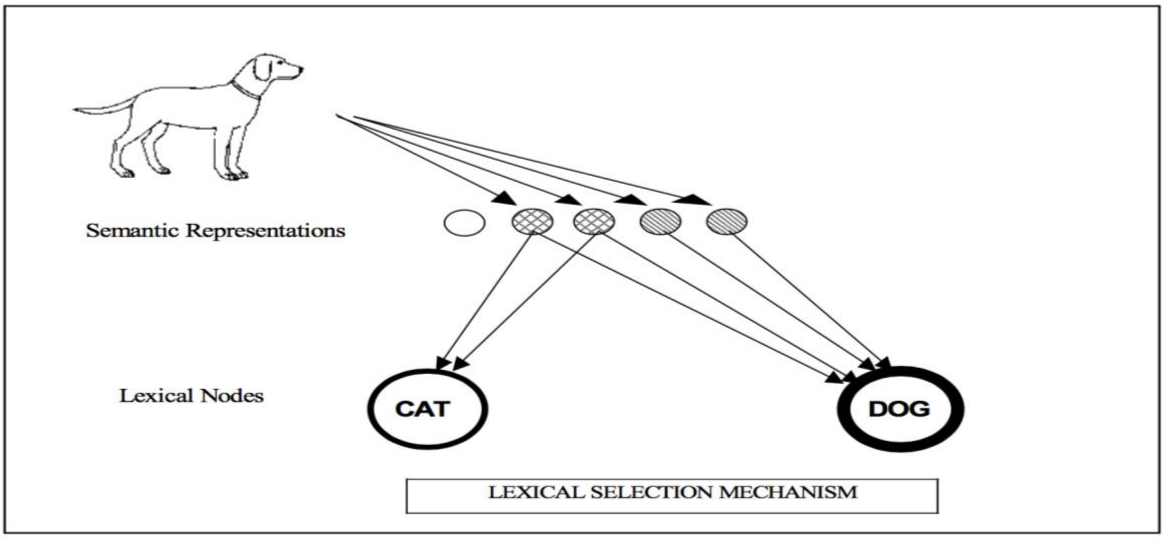


Figure. 2 Lexical access by monolingual speakers. The arrows represent the flow of activation and the thickness of the circles represents the level of activation of the lexical nodes (Costa, Colomé et al. 2000, p. 407).

Thus, a selection mechanism is required in order to identify the lexical node that corresponds to the intended concept. This mechanism chooses the lexical node with the highest level of activation and is sensitive not only to the level of activation of the target lexical nodes, but also to the level of activation of the non-target lexical nodes. That is, the higher the activation level of the non-target lexical nodes, the more difficult the lexical selection becomes. Once the target lexical node is selected, it spreads activation to the corresponding phonological representation (/d/, /ɒ/, /g/) at the phonological encoding level. The last stage of speech production involves the articulation of the target word. In general, existing theories of speech production by monolinguals recognize these characteristics of the major stages of the process. However, they differ greatly in how they are implemented (Dell 1986; Starreveld and La Heij 1995; Caramazza 1997; Roelofs et al. 1998; Levelt et al. 1999).

The next section presents a summary of the different models of monolingual speech production and their predictions regarding the flow of activation, concluding with their relevance to the extant models of bilingual lexical access.

2.1.1 Models of lexical access in monolinguals

This section describes the different monolingual models, beginning with the discrete model, followed by the cascaded model, and lastly the interactive model.

2.1.1.1 The Discrete Stage Network Model

The discrete model assumes that when a speaker names a picture of *a sheep*, for example, the conceptual representation of the target word, ‘*sheep’*, and the related conceptual representations (‘*sheep’*, ‘*milk’*, ‘*goat’*, ‘*animal’*, ‘*wool’*) are activated (Roelofs 1992). This activation then spreads to the corresponding lexical nodes at the lemma level[4], where the non-target lexical nodes act as competitors, and compete for selection (Figure 3). Thus, there is a selection mechanism, and the lexical node that receives the highest activation is selected. The activation of non-target nodes and the lexical selection mechanism are supported by the evidence of spontaneous slips of the tongue. For example, if a speaker wants to say *the dog barks,* he may say *the cat barks* instead (Costa, Colomé, et al. 2000). This type of slip is argued to be a malfunction of the lexical selection mechanism (Costa, Colomé, et al. 2000), and is termed ‘a selection error’. These slips are typically from the same grammatical class as the target (i.e., in this example, both items are nouns), and they are phonologically, morphologically, and orthographically well formed; the slip is that a non-target (semantically related item) is selected.

[4] In lexical access studies, the following terms ‘lemma level’ and ‘lexical level’ are used interchangeably, as well as the terms ‘lexeme level’ and ‘phonological level’.

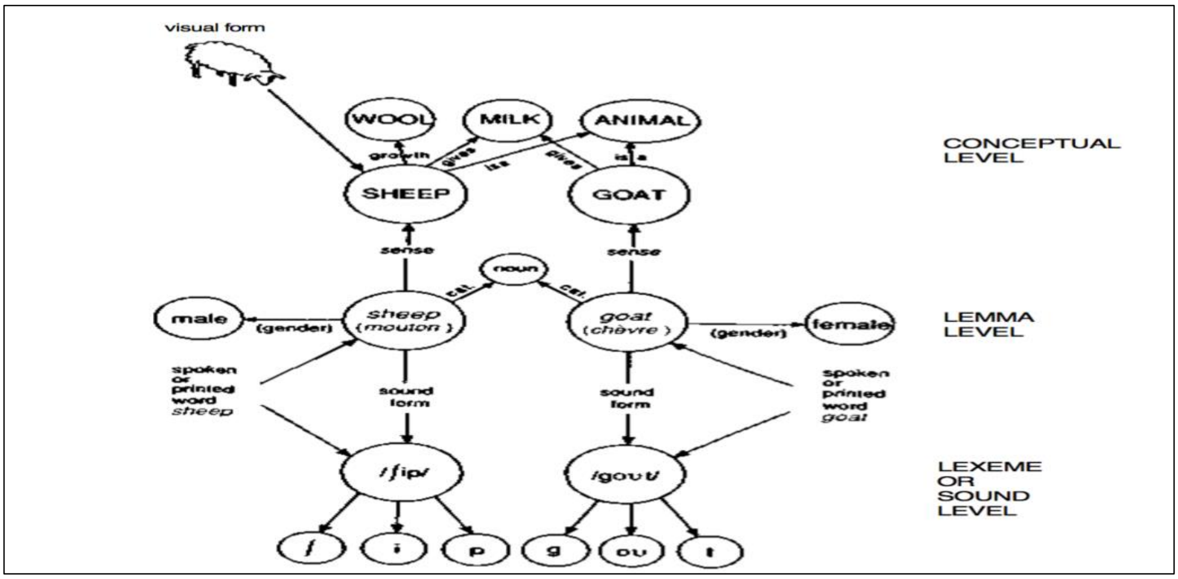


Figure. 3 A network model of lexical access (Bock and Levelt 1994, p.951).

The discrete model differs from other models in its assumption that lexical selection occurs at the lexical level, and not at the phonological level. That is, once the lexical node is selected, the corresponding phonological features are retrieved for only the selected node (Levelt 1989; Levelt et al. 1991; Roelofs 1992; Bock and Levelt 1994). The remainder of this section reviews the findings that support this view.

Evidence that supports the discrete two-stage model was provided by studies using the picture-word interference task, a variant of the Stroop task (Stroop 1935). In the study conducted by Schriefers et al. (1990), for example, the participants were asked to name pictures, while ignoring auditorily presented distractor words that were either semantically or phonologically related, or unrelated. The distractor words were presented either 150 milliseconds (ms) before the onset of the picture, 150 ms after the onset of the picture or at 0 SOA. The results showed that the semantic interference effect was obtained only in the condition in which the semantically related distractor word preceded the picture. Moreover, the phonologically related words induced a facilitation effect when presented later, namely after the onset of the picture or at 0 SOA. These findings suggested that the interference effects of the semantically related distractor words were seen early in the process of naming a picture, while the facilitation effect of the phonologically related distractor words was seen later in the process of picture naming. The researchers interpreted these results as evidence that semantic processing precedes phonological processing, which supports the discrete model of lexical access. Meanwhile, Levelt et al. (1991) conducted a similar study, in which speakers were asked to name a picture, but before doing so were presented with an auditory probe word or non-word about which they had to make a lexical decision. The relationship between the pictures and the probes was manipulated, so they were either phonologically related (for example, the probe *sheet* for the picture of *a sheep*), phonologically similar to a semantically related item (for example, the probe *goal* for the picture of *a sheep*), or semantically related (for example, the probe *goat* for the picture of a *sheep*). The timing of the presentation of the probes varied; they were presented either before the onset of the picture (an average of 73 ms), or after longer delays (373 and 673 ms). The experiment sought to prove that if semantic and phonological processing is not discrete, then when the subject names a picture of a *sheep*, the concept of *'a goat'* will be partially activated and spread activation to the lemma *'goat'*, which in turn should cause the phonological form *'goal'* to become active, much like the *'sheet'* becomes active for *'sheep'*. Thus, semantically mediated priming should be obtained. However, if the semantic and phonological processing is discrete, the partially activated lemma *goat* should not activate the phonological form of the probe word *goal*; that is, mediated priming would not be obtained. The results of the study showed that semantically related probes (*goat* for *sheep*) induced interference effects when they were presented at early SOAs, but not at later SOAs; phonologically related words induced a facilitation effect at all SOAs; and there was no effect of the semantically mediated phonological probe (such as *goal*). These findings were interpreted as evidence that semantic processing precedes phonological processing, which supported the discrete models of lexical access.

2.1.1.2 The Cascaded Model

Unlike the discrete model, the cascaded model (Figure 4) assumes that the activation of any lexical nodes (lemma) spreads activation to their phonological properties (lexeme) even before lexical selection has been achieved (Jescheniak and Schriefers 1998; Peterson and Savoy 1998). When producing the word '*dog',* for example, the activated concepts (such as '*dog', 'cat',* and *'park'*) spread activation to the linked lemma ('*dog', 'cat',* and *'park',*), which in turn activate the corresponding phonological representations. Furthermore, it is argued that the flow of activation feeds forwards, from the lexical level to the phonological level, and not vice versa.

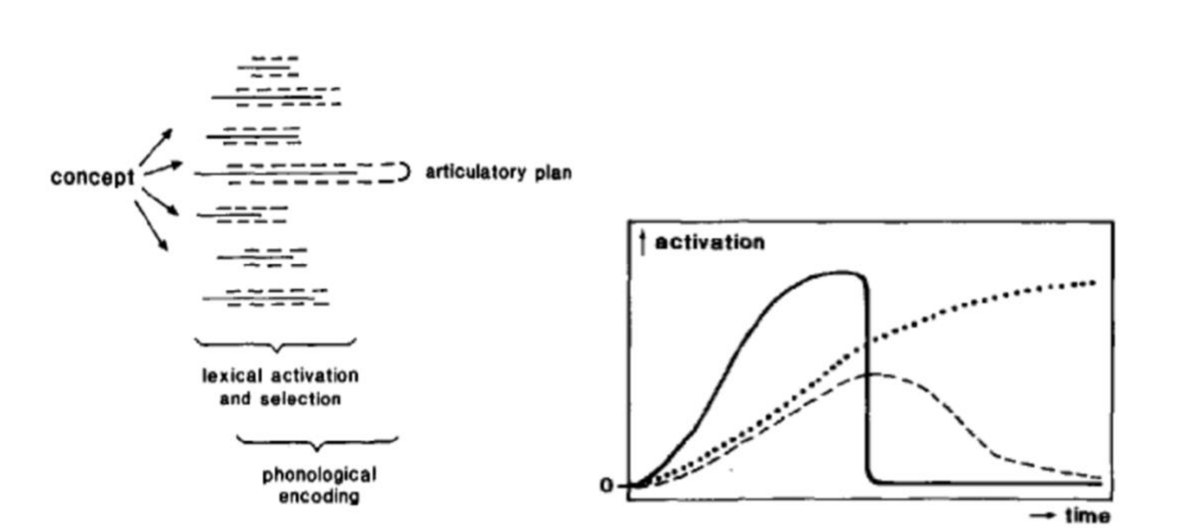


Figure. 4 A cascaded model of lexical access (Levelt et al. 1991, p.125).

Supporting evidence for the cascaded model came from previous studies that employed the picture-word interference task (Jescheniak and Schriefers 1998; Peterson and Savoy 1998). In their study, Peterson and Savoy (1998), for example, presented several experiments in which the participants had to name a picture when a question mark appeared at the offset of the picture presentation for most of the trials; whereas in the the critical trials, they were required to name a visual word presented after the picture. The relationship between the pictures and the probes was manipulated. The critical items were pictures with two synonyms (e.g., *couch* and *sofa*), words that were phonologically related to the picture names (e.g., *count* and *soda*), or words that were unrelated to the picture names (e.g., *tiger*). In this study, Peterson and Savoy (1998) hypothesized that the semantic relationship between these synonyms would induce semantically mediated phonological priming. They assumed that being interchangeable, both the lemma (e.g., '*couch'* and *'sofa'*) would be highly active during lexical access. The study proposed that if the processing is cascaded, the activation of the phonological properties of the two synonyms (e.g., *couch* and *sofa*) would be present prior to the lexical selection, resulting in the phonological properties of the phonologically related probes (e.g., *count* and *soda*) being partially activated. The results of the study demonstrated that phonologically related words were named faster than unrelated probes (e.g., *tiger*). This was interpreted as evidence for the cascaded model of lexical access, rather than the discrete model. In addition, Jescheniak and Schriefers (1998) tested the same kind of materials (in Dutch), using the picture-word interference task, and again reported results that supported the cascaded model, as phonological activation did not occur exclusively for the picture name, as the near synonym of the word was also phonologically activated. The next section reviews the interactive model.

2.1.1.3 The Interactive Model

Like the cascaded model, the interactive model assumes that the phonological activation of non-target words occurs before lexical selection. However, it differs from the other models as it assumes that this phonological activation affects the lexical selection by feeding back to any lexical nodes to which they are linked (Dell 1986; Starreveld and La Heij 1995; Dell et al.

1997). For example, when a speaker names a picture of a *dog*, before the selection of the corresponding lexical node, the phonological segments of the target word (/d/, / ɒ/, /g/) will be activated, spreading activation back to any lexical nodes containing them, such as '*doll'* and '*dot'* (Figure 5). The non-target words that are activated from the phonological system, and not from the conceptual system, act as potential candidates for selection (Costa 2004, pp. 201-223).

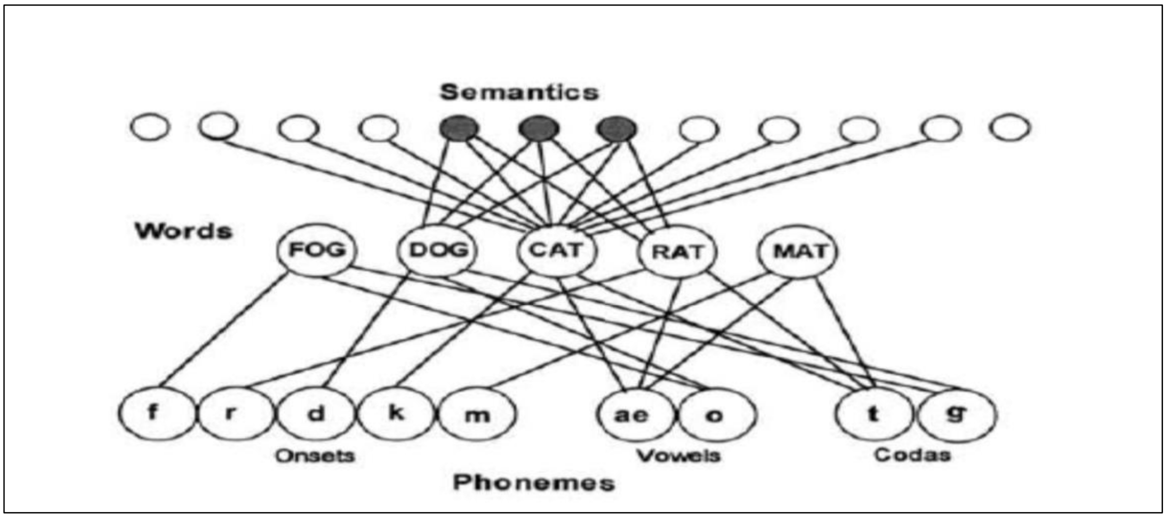


Figure. 5 An interactive model of lexical access (Dell et al. 1997, p. 805).

To illustrate, Starreveld and La Heij (1995,1996) using the picture-word interference task, asked speakers to name pictures, whilst ignoring visually presented word distractors. The relationship between the pictures and the distractor words was manipulated, so that they were either: (i) semantically and phonologically related; (ii) semantically related; (iii) phonologically related; or (iv) unrelated. The findings indicated that there was an interaction between the semantic interference effect and the phonological facilitation effect. More specifically, the semantic interference effect was reduced when the target and distractor words were both phonologically and semantically related. For example, when the participants named a picture of a *cat,* the semantic interference effect was less for the phonologically and semantically related distractor (e.g., *calf*) than for the semantically related distractor. These results suggested that activation feedbacks from the phonological level to the lexical level and affects the lexical selection process, which is consistent with the predictions of the interactive models of lexical access.

To summarize, the monolingual models recognize three different levels of processing (i.e., conceptual, lexical, and phonological), and the spreading activation principle that explains the activation flow from the conceptual to the lexical level. However, this is where the agreement ends as the models differ regarding whether they assume a discrete or cascaded flow of activation to the phonological level. Thus, research regarding lexical access by monolingual speakers primarily investigated whether phonological representations are activated at the phonological level, and if the processing is cascaded what the effect of this co-activation is on the selection process. The next section reviews how previous research concerning bilingual lexical access reported the same stages of processing, and like the monolinguals' models, how they differ regarding whether the flow of activation is discrete or cascaded.

2.2 Lexical access in speech production in bilinguals

Like the monolingual models, the current models of lexical access by bilinguals assume that there are three different stages of processing, namely conceptual, lexical, and phonological. However, the models also recognize a shared conceptual system between the two languages of a bilingual individual, and two separate lexicons (Potter et al. 1984; Kroll and Stewart 1994; Poulisse and Bongaerts 1994; Costa et al. 1999). Each conceptual representation is potentially connected to its corresponding lexical nodes in both languages. This prompts several important questions that were addressed by bilingual studies, namely: How does the spreading activation principle, from the conceptual level to the lexical level, work in bilingual lexical access? If both the target and non-target lexical nodes are activated, how does a bilingual select the target lexical node, rather than the alternative? Does the selection entail competition? And does the activation flow to the phonological level for both lexical nodes, or only to the target node?

The next section explores the spreading activation principle from the conceptual system to the lexical level, and then highlights the current disagreement in the literature regarding the lexical selection mechanism (language specific vs non-specific), and the flow of activation (discrete vs cascaded).

2.2.1 The spreading activation principle in bilingual speakers

Initially, the issue regarding the spreading activation principle in bilinguals was much debated, and two schools of thought emerged: the target language specific hypothesis[5], and the target language non-specific hypothesis (Costa, Colomé, et al. 2000). According to the target language specific hypothesis (e.g., McNamara and Kushnir 1972) the flow of activation is channelled from the conceptual system to the target language lexical system only, causing selective lexical access. It is argued that the intention to speak in the target language is sufficient to inhibit activation of the non-target lexical nodes, therefore, lexical activation is selective, and only those words in the target language receive activation from the conceptual system (Figure 6). For example, when a Spanish-English bilingual person names a picture of a *cat* in English, only the English lexical nodes/lemmas are activated.

[5] Note that we use the terms "language specific and language non-specific" throughout this study to describe lexical access and selection in bilinguals. The manner of lexical access can be language non-specific but does not entail that lexical selection is language non-specific.

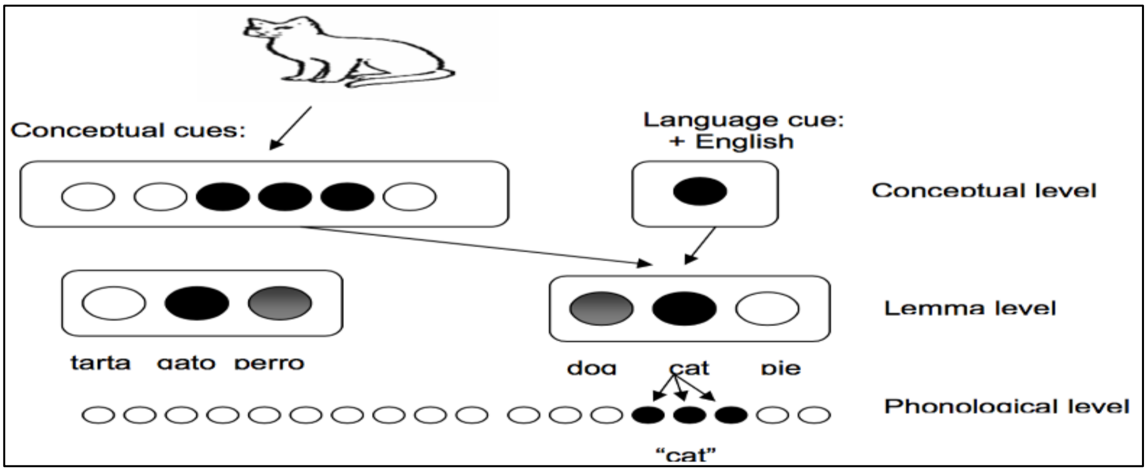


Figure. 6 A model of selective lexical activation (Hoshino 2006, p. 16).

In contrast, the target language non-specific hypothesis (Green 1986; De Bot 1992; Poulisse and Bongaerts 1994; Poulisse 1997) assumes that the activation spreads from the semantic system to both languages, regardless of the language chosen for production (i.e., non-selective lexical access). For example, when a Spanish-English bilingual person names a picture of a *cat* in English, not only are the corresponding English lexical nodes activated, but the Spanish lexical nodes are also activated (Figure 7). This view is widely accepted by the current models of bilingual word production and often referred to as non-selective activation, or the parallel activation[6] of the two lexicons (Costa and Caramazza 1999; Hermans 2000; Kroll et al. 2000). The findings of cross-language activation at the lexical level are widely reported in different tasks, such as picture-word interference tasks (e.g., Hermans et al. 1998), phoneme monitoring tasks (e.g., Colomé 2001) and language switching tasks (e.g., Meuter and Allport 1999). However, the manner of lexical selection is heavily debated. The next section discusses the different views concerning the lexical selection mechanism.

[6] Note that 'the parallel activation principle' refers to the activation of the two lexicons of a bilingual, and it is different from 'the spreading activation principle' which refers to the activation that spreads automatically from conceptual system to the other level of representations i.e., lexical/phonological level during monolinguals'/bilinguals' word production.

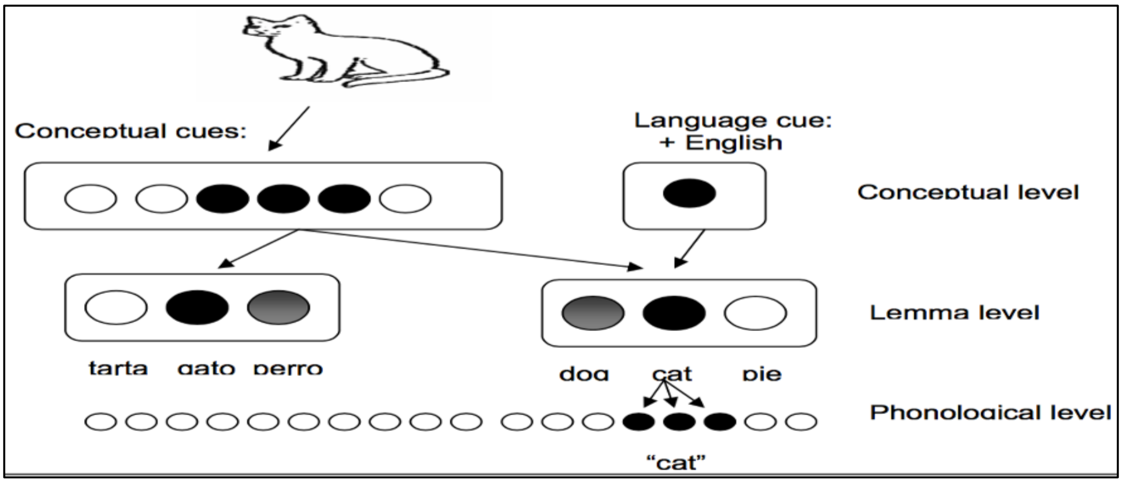


Figure. 7 A model of non-selective lexical activation (Hoshino 2006, p. 14).

2.2.2 Lexical selection mechanism: language specific or language non-specific selection?

As discussed previously in relation to monolinguals, one implication of the spreading activation principle is the activation of target and non-target lexical nodes. Hence, a selection mechanism is required to identify the target lexical node. In the case of bilinguals, since the spreading activation principle is also applied, a selection mechanism is required that not only chooses the target lexical node that corresponds to the intended concept, but is also in the correct language (Costa, Colomé, et al. 2000). As the selection mechanism is sensitive to the level of activation, it will choose the lexical nodes with the highest activation level. According to the parallel activation principle, both the target lexical node and its twin in the other language will be highly activated. So how does the speaker select the correct word, instead of its counterpart? The lexical selection mechanism can be considered in two ways: the language specific view, and language non-specific view (Figure 8).

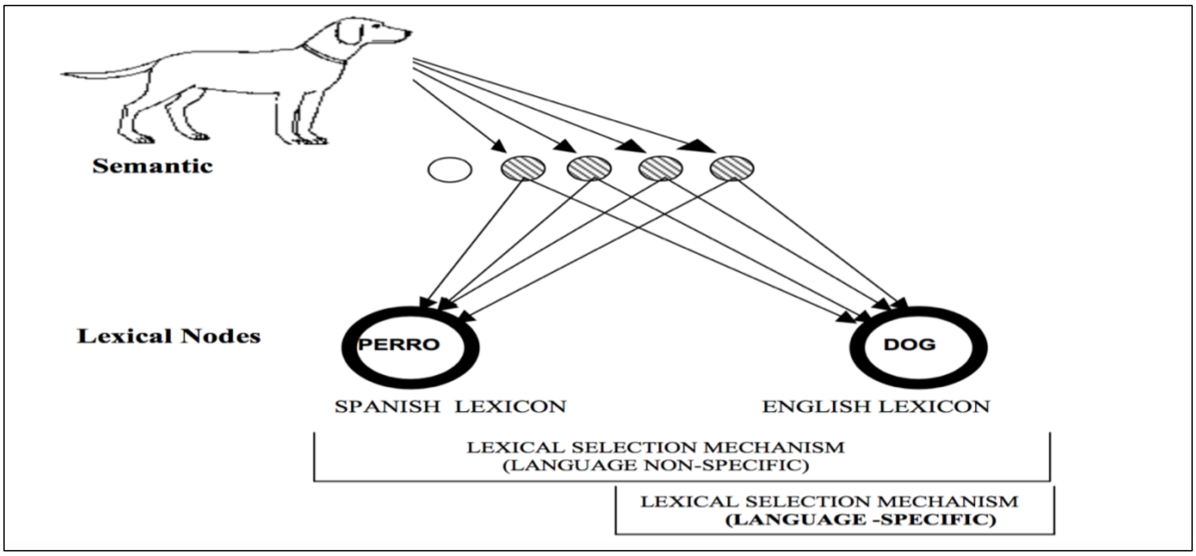


Figure. 8 Schematic representation of the language specific and non-specific selection views (Costa, Colomé et al. 2000, p. 413). The arrows show the flow of activation and the thickness of the circles represents the activation level of the lexical nodes.

The language non-specific view assumes that the lexical nodes in both languages are active and compete for selection, and there is an inhibitory mechanism (Inhibitory Control [IC]) that suppresses the activation of the lexical nodes in the non-target language (De Bot 1992; Poulisse and Bongaerts 1994; Green 1998; Hermans et al. 1998). Consequently, the activation level of the target language nodes is higher than that of the non-target language, which facilitates the selection of the former. Thus, lexical selection is language non-specific, because it considers the activation of lexical nodes in both the target and non-target language. According to Costa and Santesteban (2004), this inhibitory control mechanism is dependent on the participants' proficiency level, which is to say that unbalanced[7] bilinguals apply a language non-specific selection and rely on the IC to supress the activation of the non-target lexical node, whereas balanced bilinguals apply a language specific selection (discussed in detail later in section 2.4).

[7] The term 'balanced bilinguals' refers to bilinguals who are equally fluent in both languages and the term 'unbalanced bilinguals' refers to bilinguals who are fluent in only one language, which is usually their L1, and not fluent in the second language.

In contrast, the language specific view assumes that the lexical nodes in the target and non-target language are active, but only those in the target language are considered for lexical selection (Roelofs et al. 1998; Costa and Caramazza 1999; Costa et al. 1999). Hence, there is no cross-language lexical competition during the selection process. But how does the selection take place? Currently, only the proposal by Roelofs (1998) that there is a binding-by-checking mechanism which ensures that the word selected matches the meaning intended in the language intended has addressed this question. Further discussion of the different procedures employed in these investigations, and the contradicting findings reported in the extant literature are presented later in this chapter. The next section explores whether or not the activated target and non-target lexical nodes spread activation to the phonological level.

2.2.3 Is the flow of activation cascaded or discrete?

As with monolingual studies, the issue of whether the flow of activation from the conceptual level to the lexical level cascades to the phonological level is much debated in the bilingual studies. The cascaded view posits that any activated lexical nodes spread some activation to their corresponding phonological representations at the phonological level. This implies that the activation of the target and non-target lexical nodes cascades to the phonological level (Caramazza 1997; Costa, Caramazza, et al. 2000; Gollan and Acenas 2000), and that the selection of the target node occurs at the phonological level. In contrast, the discrete view assumes that only the lexical nodes selected are phonologically encoded (Levelt 1989; Roelofs 1992). This implies that the non-target phonological representations are not activated at the phonological level, and that the selection occurs at the lexical level. In the existing bilingual studies, there is a growing body of evidence that supports the cascaded view in similar script bilinguals (e.g., Hermans 2000; Kroll et al. 2000; Colomé 2001). However, the issue of whether the locus of selection is at the lexical level or at the phonological level remains unclear.

Previous studies assumed a single locus of selection at either the lexical level or at the phonological level. According to Costa (2005), the selection of the target lexical node must be at the lexical level, rather than at the phonological level, because the grammatical properties of the target lexical node (e.g., grammatical gender) must be accessed before the retrieval of their phonological properties. Opponents of this view employed the findings of phonological activation of non-target segments as evidence against the view of the lexical locus of selection. However, this conclusion is flawed, as finding a phonological activation of the non-target lexical nodes (i.e., cascaded activation) does not necessary entail that there is one locus of selection, and that this is at the phonological level. On the contrary, it suggests that there may be another selection mechanism at the phonological level that should be examined further. This suggestion was supported by the recent work of Blanco-Elorrieta and Caramazza (2021), who proposed the existence of a general selection process at every linguistic level. The present study explored this model further; specifically, the lexical selection mechanism was investigated in experiments one, two, and three (chapters four, five, and six, respectively), and the phonological selection mechanism in experiments one (chapter four). In experiment four and five (chapters seven, and eight, respectively), we investigated the manner of cross-language phonological activation using the phoneme monitoring task. To the best of the author's knowledge, this is the first research attempt to address the selection mechanism as a multi-level process. In addition, this study investigated whether or not script differences and participants' proficiency level modulate the flow of activation and selection.

To this point, in section 2.2., we have discussed different proposals regarding the bilingual lexical selection, the spreading activation principle, and the flow of activation. We explained that there is a consensus in the literature that assumes that activation flows from conceptual system to the lexical level, namely a parallel activation of the lexical nodes in the two languages at the lexical level, and that the flow of activation is cascaded. However, the manner of lexical selection remains unclear and is the subject of debate, namely whether or not lexical selection involves competition (i.e., language specific vs non-specific). We argued that the disagreement in the extant literature may be attributed to two factors: first, the use of different experimental paradigms; and second, the testing of bilinguals whose level of proficiency and dominance in the two languages differs. The next section critically evaluates the methodologies adopted previously, and highlights how the decision was made to focus on different levels of proficiency in the present study.

2.3 Methodologies previously adopted in bilingual word production studies

This section critically evaluates the methodologies typically employed by bilingual word production studies, namely (i) the simple picture naming task; (ii) the picture-word interference task; (iii) the phoneme monitoring task; and (iv) the code mixing/switching task. The discussion sheds light on limitations of these approaches, explaining how they might account for the contradictory findings in the literature. It also illustrates the commonly tested effects in these tasks (e.g., the cognate effect, identity effect, and the semantic interference/ facilitation effect) and their implications for models of bilingual production. It is essential that these effects are understood at this point in the thesis, as later sections of this chapter discuss how to test these effects in the masked priming paradigm, the main approach adopted by the present study.

2.3.1 Simple Picture Naming Task

Picture naming is one of the most popular paradigms for studying the processes involved in bilingual lexical access; it requires the participants to name a picture as quickly and accurately as possible, and their response time and accuracy are recorded for analysis. It is commonly used to investigate the cognate facilitation effect that refers to the advantage that cognate pictures have over non-cognate pictures in the speed of production (e.g., Hoshino and Kroll

2008). Cognates are words that are semantically and phonologically similar in the two languages (for example *gato* [Spanish, cat], *gat* [Catalan, cat]) (Costa, Caramazza, et al. 2000). It is argued that if the phonological segments of the non-target language are activated, naming latencies should be faster for cognates than for non-cognates. This is because the activation of the semantic representation of the word *'cat'* will flow to both lexical nodes, regardless of the speaker's intention to speak in one language: the Catalan lexical node '*gat'* and the Spanish lexical node '*gato'* (see Figure 9). The shared phonological segments (/g/, /a/, /t/) receive activation from both languages, resulting in a higher level of activation (Costa, Caramazza, et al. 2000) and the common finding is that the bilinguals are faster at naming pictures with cognate names than pictures with non-cognate names (e.g., Janssen 1999).

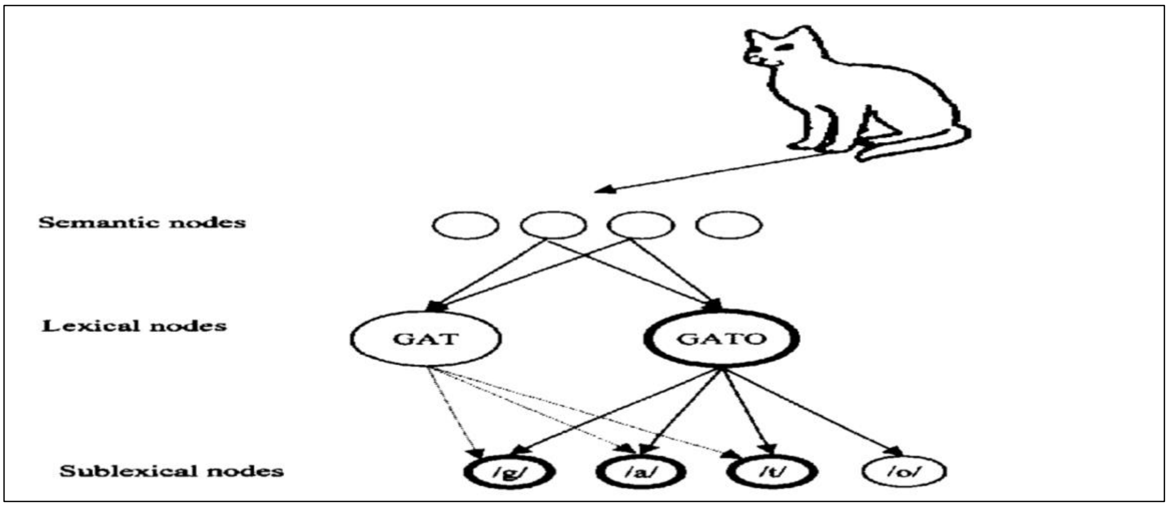


Figure. 9 Schematic representation of picture naming for cognate words (Costa, Caramazza, et al. 2000, p. 1285). The arrows demonstrate the flow of activation and the thickness of the circles shows the activation level. Some phonological segments corresponding to the Spanish target word (gato) receive some extra activation from its Catalan translation word (gat) (Costa, Caramazza, et al. 2000, p. 1285).

This differs from non-cognates, where the phonological segments receive information from one language only. The presence of the cognate facilitation effect suggests that the lexical nodes from the non-target language are activated to the point at which phonology is specified.

The results of Costa, Caramazza, et al.'s (2000) study demonstrated that the cognate effect was obtained when naming was performed in the first language (L1) and the second language (L2); however, more pronounced effects were obtained when naming was performed in the L2. The findings were replicated with a different group of bilinguals, namely Dutch-French (Janssen 1999). The cognate facilitation effect is subject to two interpretation; first, it occurs as a result of phonological overlap, as discussed earlier in this section; or secondly, that it occurs due to phonological interference caused by the activation of non-target segments at the phonological level when naming non-cognates (Costa, Caramazza, et al. 2000).

The findings of the previous cognate picture naming studies supported models of non-selective activation and language-specific selection. This is because if lexical selection entails competition, then cognates should induce interference rather than facilitation. In addition, the data suggested that both the target and non-target lexical nodes are active when phonology is specified, following a cascaded pattern of activation. In the current study, the cognate effect was examined using the masked priming paradigm to determine whether the effect is robust, even when a different procedure is applied with a different group of bilinguals, whose languages have different scripts. Another tension in this matter is that the findings concerning the faster naming of cognate pictures are often contradicted by those of picture-word interference tasks, highlighting the semantic interference effect discussed in the next section.

2.3.2 Picture-Word Interference Task

The picture-word interference task is another popular paradigm for testing the predictions of word production models. In this task, the participants are presented with pictures to name, along with visually or auditorily presented distractor words that may or may not share a relationship with them (Collina et al. 2013). The participants are usually instructed to name the pictures as quickly and accurately as possible, and to ignore the distractors. The relationship between the pictures and the distractors is manipulated, as is the language of the distractor word and the timing of its presentation relative to the pictures. The response times and accuracy are recorded for analysis. Typically, two effects are observed: the semantic interference effect, and the orthographic/phonological facilitation effect (Zhao et al. 2012). The semantic interference effect refers to the longer reaction times observed when the picture and distractor word are semantically related, for example a picture of *a dog* and the distractor word *fox*, compared to when there is no relationship between the pictures and the distractors, such as a picture of a *dog* and the distractor word *car* (Glaser and Düngelhoff 1984; La Heij 1988). Previous studies argued that the semantic interference effect reflects competition between the lexical nodes during the lexical selection at the lexical level (Schriefers et al. 1990; Roelofs 1992; Starreveld and La Heij 1995). The orthographic/phonological facilitation effect refers to the faster reaction times observed when the name of both the distractor and the picture are orthographically or phonologically related, for example a picture of a *dog* and the distractor word *doll*, compared to when there is no relationship between the picture and the distractor word (Lupker 1982; Rayner and Springer 1986; Starreveld 2000; La Heij 2005). It is argued that this effect is localized at the phonological level, as it reflects the activation of the phonological segments of the non-target lexical nodes that enhances the activation of the shared phonemes (Schriefers et al. 1990; Roelofs 1992).

For example, Hermans et al. (1998) tested Dutch-English bilinguals in a picture-word interference task in which the participants were required to name the picture in English, and to ignore the auditorily presented word in the L1. The relationship between the picture name and the distractor word was manipulated. For example, when a picture of a *mountain* was presented, the participants had to ignore the L1 distractor words that were either semantically related to the picture name (for example *dal* [Dutch, valley]), phonologically related (for example *mouw* [Dutch, sleeve]), unrelated (for example *kaars* [Dutch, candle]), or phonologically related to the Dutch translation of the picture's name *berg* (for example *berm* [Dutch, verge]). The phonologically related distractors induced shorter naming latencies, hence revealing a facilitation effect. In contrast, the semantically related distractors induced longer naming latencies and caused a semantic interference effect. Similarly, the phonologically related distractor to a Dutch translation name induced longer naming latencies, revealing the so-called 'phono-translation effect'. The semantic interference effect suggested that the lexical nodes in the non-target language were activated at the lexical level and competed for selection. The phono-translation effect occurred at the SOAs (defined in section 2.1.1.1), where semantic interference effects have been observed. Thus, this phonological interference effect was interpreted as evidence of the activation of the lexical nodes of the Dutch translation of the picture's name (*berg*) at the lexical level. In combination, the findings of the study supported a model of language non-specific selection in which the lexical nodes from both languages are active and compete for selection at the lexical level.

As stated earlier, the task has a few drawbacks, one of which is that the findings reported have multiple interpretations. For example, (Costa, Colomé, et al. 2000) argued that the phono-translation effect, found in the research conducted by Hermans et al. (1998), may not be evidence for competition at the lexical level, but could instead be interference at the phonological level. They argued that the distractor word *berm* may activate its phonological segments (/b/, /e/, /r/, /m/), some of which (/b/, /e/, /r/) may receive further activation from the lexical node '*berg'*, which is activated by the semantic representation of the picture *mountain*, and thus delay the retrieval of the phonological segments of the target word *mountain*. Moreover, Costa et al. (1999) and Costa, Colomé et al. (2000) argued that the semantic interference effect observed across languages cannot be taken as evidence of competition between lexical nodes in the target and non-target language, and thus cannot adjudicate between the language-specific and language non-specific hypothesis. According to Costa, Colomé, et al. (2000), the competition between the lexical nodes in both languages at the lexical level created by semantically related distractors in the non-target language may have two sources. First, according to the language non-specific selection view, the target word and the semantically related non-target distractor word compete for selection at the lexical level, and therefore delay the lexical selection of the target word, causing cross-language interference. Secondly, according to the language specific selection view, the semantic interference created by the semantically related non-target distractor might reflect competition between the target word and the translation of the semantically related non-target distractor word resulting in within-language lexical interference. For example, when a Spanish-English bilingual is asked to name a picture of a dog in English, the semantically related Spanish distractor *gato* (cat) activates the semantic representation of the word '*cat'*, which according to the parallel activation principle sends activation to the corresponding lexical nodes in the two languages ('*cat'* and '*gato'*) (Costa, Colomé, et al. 2000). In this scenario, the activated English lexical node '*cat'* can interfere with the selection of the target lexical node '*dog*', and thus causes a delay in the lexical selection process.

Other studies (e.g., Costa et al. 2005; Finkbeiner and Caramazza 2006; Mahon et al. 2007; Janssen et al. 2008) suggested that interference effects reflect articulatory processes outside the lexical system, that is at the articulatory output buffer, as this buffer can be engaged with only one process at a time. Moreover, they claim that related distractors prime target words at the conceptual level (Costa et al. 2005) or lexical level (Mahon et al. 2007) therefore inducing facilitative, rather than inhibitory effects. So according to these views, the interference effect is located at the phonological level and the facilitation effect is located at the conceptual level or lexical level. Abdel Rahman and Aristei (2010), in their monolingual German study, investigated the conflicting views by comparing two tasks: one that involved activating the conceptual, lexical, and phonological levels, and the other requiring only activation at the conceptual and lexical levels. The first task was a picture-word interference task in which the distractors were semantically related or unrelated to the name of the target pictures. In the second task, they asked the participants to perform a binary classification task with the name of the picture. For example, the participants were presented with a picture of a *whale* paired with a semantically related distractor word *squid* in the related condition, and an unrelated distractor word *throne* in the control condition. The participants had to manually press a button to confirm whether the last segment of the picture name was a vowel or a consonant. The same set of stimuli were used for both tasks. The results showed that in both tasks, semantically related distractors induced an interference effect, and this was interpreted as evidence upholding the lexical competition view. However, this finding cannot be considered as conclusive evidence of competition at the lexical level, as lexical representations were not the only representations activated across both tasks and, thus, the lexical level is not the only possible locus of a semantic effect. More specifically, conceptual representations were also active in both tasks, and the conceptual level may well be a better candidate for the locus of a semantic interference effect given that this is where the semantic knowledge is stored. This issue is addressed in experiment three (chapter six). The rest of this section discusses the identity effect which was employed to investigate the manner of lexical selection as a replacement for the semantic interference effect.

According to Costa et al. (1999) and Costa, Colomé, et al. (2000), the identity effect across languages can adjudicate between the language-specific selection view and the language non-specific selection view. In the identity condition, the translation equivalent of the target word acts as a distractor during picture-naming tasks. In their study, Costa and Caramazza (1999) tested Spanish-English bilinguals in a picture-word interference task. The distractor word employed was either the translation of the picture name or an unrelated word. For example, when the speaker was required to name the picture of a *dog* in English, the Spanish distractor word *perro*, the translation equivalent of the English word *dog*, was presented. If lexical selection is language non-specific, then longer naming latencies would be expected in the identity condition than in the unrelated condition. This is because the highly activated Spanish distractor lexical node '*perro'*, as it receives activation from the picture and the written word, will compete with the target language lexical node '*dog'*. However, if lexical selection is language-specific, faster naming latencies are expected in the identity condition than in the unrelated condition. This is because the target lexical node receives extra activation from the Spanish distractor word through its semantic representation and because, according to the language-specific view, only target language lexical nodes are considered for selection, and the activated non-target language lexical nodes are ignored. The study found that faster naming latencies were observed when the distractor word was a translation equivalent of the target word, regardless of the language in which the distractor word was printed. These results were in accordance with the predictions of the language-specific view, which argues that the lexical selection mechanism considers only the activated lexical nodes in the target language. The same results were obtained when Costa et al. (1999) tested Catalan-Spanish bilinguals. Faster naming latencies were obtained in the identity condition than in the unrelated condition and the authors thus concluded that the identity effect phenomenon is robust. However, it could be argued that the processing mechanism within the two languages is similar, given the phonological/orthographic similarities between Spanish-English, and so the faster naming latencies might be due to this overlap. It could also be attributed to the use of distractor words that share the same onset as the picture name. For example, the participants were presented with a picture of *a nose* with the Spanish distractor word *nariz* (the translation of the target English word *nose*). The evidence would be more conclusive if the same finding was obtained when testing bilinguals with languages that employ different scripts, such as the Arabic and

English languages using the same/different paradigm, and this was one of the gaps addressed by the current research.

As discussed previously, the picture-word interference task is frequently used to understand the bilingual production process, and specifically to explore whether lexical nodes are activated in the non-target language while producing the target lexical nodes in the target language. Previous researchers argued that this task is not ideally suited to answering this question, because it is unclear whether this parallel activation occurs due to top-down activation (i.e., naming a picture which starts by conceptual activation) or bottom-up activation (i.e., from the presentation of the printed distractor word) (Costa, La Heij, et al. 2006). According to Kroll et al. (2010, p. 3)

> in word recognition, there is evidence for bottom-up parallel activation of word form information in both languages (e.g., orthography and/or phonology). In word production, there is evidence of top-down activation of meaning-related neighbours (e.g., semantic relatives in both languages, including translations).

Therefore, this study employed the masked priming paradigm in picture naming that enabled the elimination of word recognition processing (i.e., reading the visually presented word) by applying a mask over the prime word. In this way, the effects of the picture word interference task reported in the previous literature could be reinvestigated in a similar paradigm, whilst avoiding its drawbacks. The priming paradigm was not the only method of investigation employed in this study, as the phoneme monitoring task was also used to test lexical access and the flow of activation in different script bilinguals. Thus, the next section provides a comprehensive review of these methods, including the most common findings reported in the extent literature, and their interpretations.

2.3.3 Phoneme Monitoring Task

In previous studies, the phoneme monitoring task was used to explore the phonological representations involved in speech production in bilinguals (Wheeldon and Levelt 1995; Colomé 2001). In this task, participants listen to an auditorily presented word and are required to decide whether a specific phoneme, or a letter corresponding to that phoneme, is present. For example, Wheeldon and Levelt (1995) asked Dutch-English bilinguals to listen to a list of English words and to translate the English words into Dutch in the first task. Then, in the second task, they were presented auditorily with a phoneme followed by an English name. The participants had to press a button when the Dutch translation of the English name contained the target phoneme. Later, Colomé (2001) adopted this task to investigate whether non-target phonological representation is active during word production. The study required Catalan-Spanish bilinguals to decide whether or not a Catalan target phoneme was present in Catalan picture names. The presentation of the printed target phoneme preceded the presentation of the picture. For example, in the critical condition, the participants were presented with a picture of a *taula* (the Catalan word for table) and asked whether the phoneme /m/ (which is present in *mesa*, the Spanish name for table) was present in the Catalan name of the picture. In the unrelated condition, they were asked whether the phoneme /f/ was present in the Catalan name of the picture. In the study, Colomé (2001) argued that if the non-target Spanish lexical node '*mesa'* is active, along with its phonological segments, longer reaction latencies would be expected when rejecting the phoneme /m/, than when rejecting the phoneme /f/. The study's results showed that the reaction latencies were longer when the target phoneme was part of the non-target Spanish word than when rejecting the unrelated phonemes. These findings were interpreted as evidence for the cascaded model. When a picture of the table is presented, its conceptual representation activates the lexical nodes in the target and non-target language ('*taula'* and '*mesa'*), which in turn sends activation to their phonological segments at the phonological level. Moreover, Hermans (2000) obtained the same results when testing bilinguals with a different language pair (Dutch-English). However, the phoneme monitoring task has some disadvantages. First, it is possible that the presentation of a non-target phoneme might trigger the activation of the non-target name of the picture. Also, these studies employed languages with similar scripts (e.g., Dutch-English), and phonologically similar languages (e.g., Catalan-Spanish), which might contribute to the activation of the non-target language. Meanwhile, Hermans et al. (2011) argued that Colomé (2001) used a list of stimuli that consisted of cognate and non-cognate filler items, which would account for the cross language phonological activation found in that study. Therefore, Hermans et al. (2011) employed the same task to determine whether the magnitude of phonological activation of the L1 names increases if cognate pictures are used as fillers in the task. They conducted three experiments in which Dutch-English bilinguals were asked to decide whether a specific phoneme was part of the L2 name of the pictures concerned. The size of the cognate fillers was manipulated in the three experiments; the list of picture names had 0 cognate names in experiment one, 100% cognate names in experiment two, and 25% in the third experiment. Cross language phonological activation was found only in experiments two and three, but not in one. The authors concluded that the findings suggested a dynamic production system in bilinguals that can operate in different modes, depending on the composition of the stimuli list. However, despite these drawbacks, the phoneme monitoring is the only task available that can access cross-language phonological activation, especially the unshared phonemes across the two languages. Thus, this study adopted the task to investigate whether non-target phonemes (shared and non-shared) are active during the production process of different script bilinguals.

In short, previous studies employed the phoneme monitoring task to investigate whether the activation of the target and non-target languages cascades to the phonological level. Most of the results supported a cascaded model, in which the lexical activation is language non-specific, and the activated lexical nodes of both languages are specified at the phonological level. The next section reviews the final method, code switching/mixing, which is the most widely-employed paradigm when investigating cross-language competition, and how it is resolved by an inhibitory control mechanism, as discussed in section 2.2.2. Most previous code mixing/switching studies explored the role of proficiency level, and whether it modulates lexical selection. Since the current study addressed the same question, it is important to review their findings to ground the present research appropriately.

2.3.4 Code Switching/Mixing

The phenomenon of code switching is the ability of the bilingual speaker to switch from one language to another, often several times in a single utterance (Kroll et al. 2012). As Kroll et al. (2012, p. 232) explained, "The observation that both languages are active but that bilinguals are able to select the intended language with relative accuracy suggests that they develop cognitive control that enables them to negotiate the potential cross-language competition". Several studies employed the code switching/mixing paradigm to investigate whether lexical access in bilingual word production involves cross-language competition, and the possibility of inhibitory control. For example, Miller (2011) asked bilinguals to name pictures, or to read words/numbers in one of their languages in a mixed language sequence and reported that language mixing affects the L1 and the L2 differently which supports the non-specific selection view. Similarly, Meuter and Allport (1999) asked unbalanced bilinguals to name numerals in their L1 and L2 unpredictably as they had to switch language according to the background colour on which each numeral was displayed. For example, the English-French participants were instructed to name the number in their L1 if the background was blue, and in their L2 if the background was yellow. The results demonstrated that the response latencies in the switch trials were slower than those in the non-switch trials indicating that the cost of language-switching[8] was higher when switching from the less dominant language (L2) to the dominant language (L1) than when switching from the L1 to the L2, as the L2 responses were faster than the L1 responses, showing that there was greater inhibition of the dominant language (L1) during the production of the weaker language (L2). The greater suppression of the L1 during the production of the L2 made it harder to reactivate the L1 in subsequent trials, in which the L1 was the target language. The findings support a model of non-selective activation and language non-specific selection, in which the two languages are active at the lemma level and compete for selection.

In addition, the suppression of the L1 observed in language switching tasks was in accordance with the assumptions of the Inhibitory Control Model proposed by Green (1998) which assumes that both the L1 and the L2 are active at the lexical level (both are activated from the semantic system), and that the selection mechanism is sensitive to the level of activation of the target and non-target lexical nodes. The level of activation of the non-target lexical nodes is controlled by a task schema. It is assumed that the task schema suppresses the activation of the non-target language when the speaker intends to speak in the other language. Moreover, the model assumes that it is more difficult to inhibit the activation of the L1 than the L2, because the former is normally more active. Several previous studies have reported a similar pattern of results, with larger switch costs from the weaker language to the dominant language (e.g., Jackson et al. 2001; Philipp et al. 2007; Schwieter and Sunderman 2008). However, the participants in these studies were unbalanced bilinguals; that is, one of their languages was dominant. In order to explore whether this inhibitory control is also present in the language production of balanced bilinguals, or if it is exclusive to unbalanced bilinguals, Costa and

[8] The term ‘switch cost’ refers to the phenomenon that bilinguals have worse performance in switch trials relative to non-switch trials.

Santesteban (2004) tested two groups of unbalanced bilinguals (Spanish-Catalan bilinguals and Korean-Spanish bilinguals) and a group of highly proficient balanced bilinguals (Spanish-Catalan bilinguals). The study's participants were asked to name pictures in a language-switching task. They were informed that the language in which they were required to name the picture was determined by the colour in which the picture appeared. For the two groups of unbalanced bilinguals, the results showed that switching from the non-dominant language (L2) to the dominant language (L1) was harder than vice-versa. In contrast, the balanced bilinguals produced symmetrical switch costs when switching between their two dominant languages. According to Costa and Santesteban (2004), this symmetrical switch cost was due to the fact that the balanced bilinguals were equally proficient in their two languages (L1 and L2), and the amount of suppression needed to speak in the intended language was thus similar for both languages. An unexpected finding of the study was that the balanced bilinguals produced the same symmetrical switch costs when switching between their dominant L1 language and their much weaker third language (L3), namely English. Accordingly, Costa and Santesteban (2004) argued that once they have reached a high level of proficiency in at least one of their additional languages, balanced bilinguals do not use inhibitory control, and instead apply a language-specific selection mechanism, even for their less proficient L3. Moreover, Calabria et al. (2012) replicated the same symmetrical findings with highly proficient Catalan-Spanish bilinguals. The results of these studies suggest that the presence of inhibition in bilingual language production depends on the level of proficiency of the speakers (see section 2.4 for further discussion), a claim challenged by the present research.

Costa, La Heij, et al. (2006) argued that whilst these language-switching studies might be helpful for understanding the control mechanisms used by bilingual speakers during word production, they are not informative regarding the question of whether the non-response language is active during the speech production process; "This is because, arguably, in a language switching task participants may have their two languages active in a way that is not comparable to cases in which they are speaking in only one language" (Costa, La Heij, et al. 2006, p. 141). This gap in understanding was addressed in the present study through its examination of two proficiency groups in a monolingual masked priming task in which any factor that might trigger the activation of the non-target language deliberately were eliminated.

Kroll et al. (2000) developed a cued picture naming paradigm, in which they combined the picture naming and the language-mixing paradigms. Highly proficient Dutch-English and French-English bilinguals were instructed to name pictures in one of their two languages after the presentation of a tone cue. In the mixed language condition, a high tone was used to signal one language and a low tone to signal the other language. In the blocked condition, a tone was used to signal constant naming in one language. The cued picture naming task sought to compare the performance of the bilinguals in the mixed condition, in which both languages were deliberately forced to be active, and in a blocked condition, where only one language was required to be active. If the L1 was active during the production of the L2, then forcing it to be active should have few consequences for performance. To investigate whether phonological information relating to the non-target language was active at the phonological level, the naming latencies of the cognate and non-cognate names of the pictures were compared. If the lexical nodes of the target and non-target languages were active at the phonological level, the cognate facilitation effect should have been observed. The study found that the facilitation effect was obtained in the mixed condition for both the L1 and the L2, but only for the L2 in the blocked condition. The participants were slower at naming the pictures in the L1 in the blocked condition than in the mixed condition, whereas there was no difference between the mixed and blocked conditions in naming the pictures in the L2. In other words, requiring both languages to be active came at considerable cost to the L1, but not to the L2. The results were interpreted as evidence of the normal activation of the L1 during L2 production, and of the activation feeding forward to the phonological level. The data supported a model of non-selective lexical access, in which the target and non-target lexical nodes are activated at the lexical level, and cascade to the phonological level. Kroll et al. (2000) challenged the cognate effect in the language mixing task and tested highly proficient bilinguals whose languages shared the same script, and the present study, sought to extend this research to include different script bilinguals, employing a different paradigm.

In summary, among the extant literature there is consensus regarding non-selective lexical access, namely the parallel activation of the two languages during bilingual word production. However, the manner of lexical selection remains unclear, and the flow of activation is viewed recently as cascaded in similar script studies. Furthermore, the question of whether proficiency level affects language activation and the manner of lexical selection was rarely explored. An important point to note is that the findings reported previously were from studies that involved bilinguals who spoke languages that employed the same script, namely the Roman alphabet. A shared script may contribute to the finding of non-selective lexical access observed by these studies. It is possible that different script would act as a language cue and direct bilinguals to selectively access the target language (further discussion of this view is in section 2.6). To the best of the author's knowledge, only a few studies investigated the selectivity issue among bilinguals whose two languages have different scripts, such as Japanese-English (Hoshino 2006) and Korean-English (Moon and Jiang 2012). These studies adopted the same experimental paradigms used in testing same-script bilinguals, namely simple picture naming, picture-word interference tasks, language switching, and phoneme monitoring tasks which have some drawbacks as we discussed in this review in sections 2.3.1, 2.3.2, 2.3.3, and 2.3.4. The next section reviews the previous studies that involved different script bilinguals, and concludes with the justification for the use of the chosen procedures for the present study.

2.3.5 Lexical access in different-script bilinguals

As discussed previously, compelling evidence exists that shows that lexical access is non-selective in bilinguals whose languages have similar scripts. In order to establish whether this finding can be generalized to bilinguals whose languages have distinct scripts, a small number of studies examined lexical access in different script bilinguals (e.g., Kheder and Kaan 2019); two of these are reviewed in this section.

Hoshino (2006) compared the performance of bilinguals whose languages possessed different scripts (Japanese-English) and bilinguals whose languages shared the same script (Spanish-English) in two different tasks: picture naming and picture-word interference. In the L2 picture naming task, the performance of the bilinguals in naming pictures with cognate and non-cognate names was compared and cognate facilitation effect was observed in the performance of both groups of bilinguals supporting the cascaded view that both languages are activated at the level of phonology. In the L2 picture word interference task, four types of L1 distractor words were used: semantically related, phonologically related, translation, and phono-translation distractor words. The aim of the picture-word interference task was to determine whether the script differences modulated the cross-language activation and the locus of selection, since the written script was present in this task, unlike in the simple picture naming task. The language in which the distractor word was presented was manipulated, as was the timing of the presentation of the distractor word relative to the pictures, and the relationship between the picture and the distractor words. The key finding was that both groups of bilinguals exhibited the phonological and translation facilitation effect, whereas only the Spanish-English bilinguals demonstrated both semantic interference and the phono-translation facilitation effect.

The presence of the phonological facilitation effect in the absence of semantic interference for the different script bilinguals was explained as being due to the recognition of the L1 distractor being automatic, which caused the immediate activation of the L1 phonology, and influenced the speech planning in the L2. By the time the semantic representation of the L1 distractor became active, the different script bilinguals exploited the perceptual information of the unique scripts of the distractor word as a language cue, thus inhibiting the activation of the L1 semantic representation, and selecting the language of production. Thus, only the target lexical nodes were considered for selection. The results supported the fact that two processes are involved in picture-word interference, namely word reading and word production. In word production, the semantic representation is activated first, whereas in word reading the phonological information is activated before the semantic representation. Thus, the presence of phonological information prior to semantic information is expected in the picture-word interference task. It could therefore be argued that the picture word interference task does not tap into the process of speech production exclusively.

The presence of the translation facilitation effect, and the absence of the phono-translation facilitation effect was also explained by the time course during processing. The L1 translation distractor word sends activation to its semantic representation. This extra activation at the conceptual level, from the picture and the translation distractor word among other related concepts, contributed to the facilitation effect found. In contrast, in the case of the phono-translation distractor, the process takes longer, as the L1 distractor is required to activate its phonological representation first, which then causes the activation of the L1 translation name of the picture, which in turns sends activation to the conceptual level. By the time the conceptual representations of the L1 distractor word were activated, the Japanese-English bilinguals may have selected the target lexical node. This finding was interpreted as evidence that the distinctive script of the distractor words facilitates language and lexical selection. This finding suggested that the manner of lexical selection in different script bilinguals is modulated. This view was challenged by the current study, as the identity effect (i.e., translation facilitation effect) was examined, along with the semantic interference effect, with a different group of bilinguals whose languages have distinct scripts.

Combining the findings, Hoshino (2006) concluded that: (i) lexical access is language non-specific, and that different scripts facilitate language selection at an early point of production; and (ii) the manner of lexical selection in different script bilinguals is inconclusive, as it is unclear whether these bilinguals attend to the target language selectively, or supress the non-target language earlier in the process. This matter was addressed further by the present research (see chapters four, five, and six).

Moon and Jiang (2012) also investigated whether lexical access is non-selective in bilingual speakers whose languages have different scripts. They adapted Colomé’s (2001) phoneme monitoring task to test highly proficient Korean-English bilinguals. The participants were presented with pictures and asked whether the name of the picture contained a target phoneme. The target phoneme was visually presented, followed by the picture. The task was conducted in Korean and in English. For example, in the Korean task, there were three conditions: (i) a positive condition, in which the Korean picture name contained the target phoneme; (ii) a critical condition/ interference condition, in which the target phoneme existed in only the English picture name; and (iii) a control condition, in which the target phoneme was not present in either the Korean or the English picture name. The rationale of the task was that if the non-target language was active, then longer response latencies would be observed in the interference condition, where the target phoneme existed in the name of the picture in the non-target language, than those in the unrelated condition. If the non-target language was not active, no difference should be found between the interference condition and the unrelated condition. The results showed that longer response latencies were observed in the interference condition than in the unrelated condition. This was reported as evidence that the non-target language is activated during the production of the target language, and that the response delay is because of competition between the activated lexical nodes from the two languages. In addition, it was claimed that language dominance does not affect the non-selective activation of the two languages. Thus, the authors concluded that lexical access in bilinguals whose languages have distinct scripts is non-selective and this finding will be compared with the results here from Arabic-English (different script) bilinguals.

In summary, the findings of these two studies suggested that lexical access in different-script bilinguals is non-selective, and that script differences cannot direct lexical access selectively (Hoshino 2006; Moon and Jiang 2012). However, the differences trigger bilinguals to select the language of production at an earlier point in speech planning (Hoshino 2006). It should be noted that there are limitations to Hoshino's (2006) experimental paradigms, as mentioned earlier. Specifically, it was argued previously that the picture-word interference task is not ideally suited to investigating the parallel activation of two languages during the course of production, as it is unclear whether this parallel activation occurs because of top-down activation (i.e., naming a picture) or bottom-up activation (i.e., from the presentation of a printed distractor word) (Costa, La Heij, et al. 2006). In addition, the effect of proficiency level on lexical access has not been tested adequately to date, as previous studies tested only highly proficient bilinguals (e.g., Moon and Jiang 2012), and thus, the present research also tested the abilities of less proficient, different script bilinguals to determine whether the results illuminated the nature of lexical access further.

The present study employed the masked priming task to test the reported effects found in the picture word interference tasks, namely the cognate facilitation effect, the translation facilitation effect, and the semantic interference effect, as these were the major effects reported previously that shaped the current models of word production. To the best of the author's knowledge, this masked priming paradigm in picture naming has rarely been used in previous studies that tested lexical access in bilinguals whose languages have a shared script. We note a single Event-Related Potentials (ERP[9]) study which used this paradigm to study English-French bilinguals' picture naming processing (Chauncey et al. 2009) (for a review of ERP method see Blackwood and Muir 1990). The study did not assess the flow of activation and whether it was cascaded or discrete, or the manner of lexical selection in bilinguals. In terms of different script bilinguals, this paradigm has never been used in the investigation of lexical access in general, or in exploring the manner of lexical selection and the flow of activation.

The present study tested Arabic-English bilinguals in a masked priming task; this group has rarely been investigated in bilingual lexical access studies. In addition, this study also employed the phoneme monitoring task to test the effect of the different scripts on cross language-activation and the flow of activation, and to explore whether the differences in proficiency level affected lexical access, cross-language activation, and the manner of lexical selection.

The next section discusses the relevance of studying word production in different script bilinguals, exploring the role of proficiency level and script differences in the cross-language activation. It then introduces the cross-linguistic differences between Arabic and English languages.

[9] ERPs are "very small voltages generated in the brain structures in response to specific events or stimuli" (Blackwood and Muir 1990, p. 96). In psycholinguistics research, this method is used to record the brain waves when it processes certain event such as speech production, perception, etc.

2.4 The role of proficiency level in bilingual word production

Before discussing the role of proficiency level in the production process, it is important to understand what language proficiency level is, and how researchers typically assess an individual's level of proficiency in any language. Language proficiency level refers to how well an individual uses language in its oral and written form in different situations (Cloud et al. 2000). Previous researchers employed different assessment methods to examine bilinguals' proficiency level, either for controlling purposes, namely to ensure that all of their participants were matched on proficiency level, or for categorizing them under different proficiency groups (e.g., highly proficient group vs less proficient group). These methods fall into three categories: (i) the standardized tests, such as the International English Language Testing System (IELTS) test; (ii) the placement test administered by teaching units in educational institutions to place students into language classes that match their level; and (iii) proficiency measures/tests administered by academic researchers, such as lexical decision tasks, and self-rating in the language history questionnaire. In word production studies, the most frequently used measure is self-rating in the language history questionnaire (e.g., Miller and Kroll 2002), which is sometimes is combined with another assessment tool, such as the lexical decision task (e.g., Hermans 2004; Jacobs et al. 2016). We believe it is important to combine two or three assessment methods of language proficiency testing to obtain robust categorizations. Depending solely on self-rating, for example, may cause the inaccurate classification of proficiency, because some participants might overrate/underestimate their proficiency level. Also, placement tests can differ greatly across institutions, as highly proficient bilinguals at one institution, might be classified as less/intermediate proficient bilinguals at a different institution, unless they are all rated according to a single unified test. Thus, this study (cf chapter three) used more than one assessment measure to ensure that the participants were categorized robustly as either highly or less proficient. This section now reviews the published work concerning the role of proficiency level on word production processes, and whether the findings were conclusive.

Despite the growing interest in understanding the role of proficiency level on word production processes (e.g., Kheder and Kaan 2019), it remains unclear whether proficiency level modulates the lexical selection and retrieval process. Nevertheless, there is compelling evidence that highly and less proficient bilinguals experience parallel activation of their two languages (e.g., Kheder and Kaan 2019), and that as proficiency increases, the speed and accuracy of lexical retrieval in the L2 are enhanced (Costa and Caramazza 1999; Kroll et al. 2005). However, the extant findings regarding the manner of lexical selection in highly and less proficient bilinguals are contradictory. While Costa, Santesteban, et al. (2006) argued that less proficient bilinguals rely on language non-specific lexical selection (defined in section 2.2.2), and highly proficient bilinguals employ a language specific manner, Hermans et al. (1998) argued that the manner of lexical selection was language non-specific for all the Dutch-English bilinguals in their study. Moreover, others (e.g., Kroll et al. 2006; Hermans et al. 2011; Grosjean and Li 2013; Boukadi et al. 2015) claimed that the manner of lexical selection is dynamic, and is dependent not only on proficiency level, but also on other factors, such as language context and semantic constraints. For example, Kheder and Kaan (2019) investigated lexical selection, cross language interaction, and switch costs in Arabic-French bilinguals. In their study, the participants listened to a sentence in French or Arabic (e.g., I need money, I have to go today to [..….]), and then performed a naming task on a visually presented target word (e.g., *the bank*) that completed the sentence they had heard. In the switch trials, the sentences were in Arabic and were completed with a French target word, and in the non-switch trials the sentences and the target word were all in French. The study manipulated two factors: the sentence context, namely semantically constraining towards the target word (e.g., every time we brush the teeth, we should rinse [*the mouth*]), and neutral, (e.g., this boy did not sleep, because he had pain in [*the mouth*]), and the type of target words, namely cognates (e.g., *the bank*) and non-cognates (e.g., *the mouth*). The findings demonstrated that there was significant interaction between the cognate effect, and the semantic and language contexts in the less proficient group only. Based on this finding, Kheder and Kaan (2019) concluded that lexical selection is more language specific for highly proficient bilinguals, but more language non-specific for less proficient bilinguals, supporting the claim that the stronger the L2, the more language specific lexical selection is.

A possible reason for the contradictory results is that the conclusions were drawn from different studies that employed different procedures, applied different proficiency measures, and compared the effect of proficiency level within bilingual/trilingual groups (i.e., comparing between the L1 and L2, or between the L1 and L3). For example, the findings of Costa, Santesteban, et al. (2006) were based on trilinguals' performance in their L3 and their L2. Meanwhile, the study conducted by Hermans et al. (1998) featured bilinguals whose level of proficiency was ambiguous, as it was recorded that the participants had received five years of education in English, but it was not known whether they were beginner, intermediate, or advanced L2 learners.

Most existing studies in bilingual word production tested mainly balanced bilinguals, and currently only a small number of studies compared the performance of two different proficiency groups in an L2 (e.g., Kheder and Kaan 2019), and these were mostly in language switching tasks, as reviewed earlier in this chapter. In order to establish the role of proficiency level in bilinguals' word production, further investigation is needed of less and highly proficient bilinguals that conducts a comparison between the bilinguals' performance in their L2. The present research sought to fill this gap in understanding, and to untangle some of the contradictory findings. It tested two different proficiency groups in their L2 using tasks not typically employed in word production studies, namely a masked primed picture naming task, an animacy decision task, and a phoneme monitoring task. Moreover, the bilinguals recruited in the previous studies were mostly same script bilinguals. Only recently have a limited number of studies sought to bridge this gap by conducting investigations with different script bilinguals (e.g., Moon and Jiang 2012; Boukadi et al. 2015); the present research added to this by working with Arabic-English (different script) bilinguals.

The differences between highly and less proficient bilinguals in terms of lexical retrieval and selection in production was evaluated previously in relation to two theories: the Language Inhibitory Control theory (IC) (Green 1998), which was introduced earlier in this chapter (2.2.2), and the Revised Hierarchical Model (RHM) (Kroll and Stewart 1994). Thus, it is important to review these theories and their implications, as they are utilized in the discussion of the findings of the present study.

2.4.1 Language Inhibitory Control

As discussed earlier in this chapter (section 2.2.2), the language non-specific view proposes that competition between the lexical nodes in the two languages is resolved through the inhibitory control mechanism that supresses the activation of the non-target lexical nodes (De Bot 1992; Poulisse and Bongaerts 1994; Green 1998; Hermans et al. 1998). According to the IC model, the dominant language is more strongly inhibited than the weaker language among unbalanced bilinguals. So, for less proficient bilinguals, stronger inhibition is applied to the L1 than the L2, whereas for balanced bilinguals, symmetrical inhibition is applied to both the L1 and the L2. Therefore, previous studies that employed language switching tasks (reviewed in section 2.3.4) reported that highly proficient bilinguals exhibit symmetrical switching costs (namely, reactivation costs), whereas less proficient bilinguals exhibit asymmetrical switching costs (Costa and Santesteban 2004; Costa, Santesteban, et al. 2006), indicating that as proficiency increases, the language control mechanism is enhanced during language processing (Meuter and Allport 1999; Costa and Santesteban 2004; Costa, Santesteban, et al. 2006; Blumenfeld and Marian 2007; Verhoef et al. 2010; Mosca and de Bot 2017). In brief, less proficient bilinguals apply language non-specific selection, and thus use inhibitory control to supress the activation of the non-target language, while highly proficient bilinguals apply language specific selection, and thus do not need to use the inhibitory control mechanism (Costa and Santesteban 2004). The present study tested the predictions of this model in experiments one, two, and three (chapters four, five, and six, respectively) as the main focus of these experiments was lexical selection in bilinguals.

2.4.2 Revised Hierarchical Model (RHM)

The RHM concerns the nature of the connections between concepts and words in bilinguals' two languages (Kroll and Sunderman 2003), assuming that there is an asymmetry in the strength of the connections between the words in the two languages and concepts (Figure 10). The model accounts for the development of conceptual processing with increasing L2 proficiency level, and suggests that less proficient bilinguals have weaker links between their L2 lexical nodes and their conceptual representations, whereas L1 lexical nodes are strongly connected to their conceptual representations, and have direct access (Kroll and Stewart 1994).

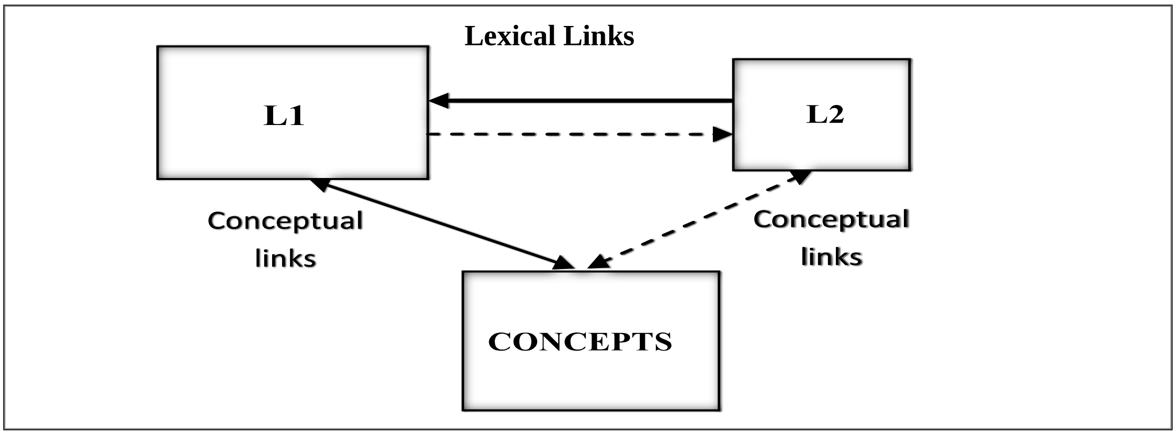


Figure. 10 The Revised Hierarchical Model (Kroll and Stewart 1994, p. 158). The dashed lines signal weak links, and the solid lines signal strong links. The arrow signals the direction of the access.

The weak links between an L2 and the conceptual system are asymmetrical in the sense that access from lexical nodes to their conceptual representations is accomplished easily, but access from conceptual representations to the corresponding lexical nodes is effortful (Kroll et al. 2010). These weak links reflect negatively on their performance, causing delayed L2 lexical activation and retrieval compared to that of the L1 (e.g., Van Hell and Tanner 2012). In contrast, highly proficient bilinguals have symmetrically strong links between L2 lexical nodes and their conceptual representations, thus faster retrieval and activation of L2 is expected than in less proficient bilinguals.

In summary, these two theories explain why highly proficient bilinguals are faster at lexical retrieval than less proficient bilinguals, and explain that they apply language specific selection, whereas less proficient bilinguals apply language non-specific selection. These two theories are complementary in the sense that one accommodates for the contradictory findings in the literature regarding the manner of language selection, and the other one accommodates for the findings of enhanced lexical retrieval as proficiency level increases. The current study tested these assumptions on different script bilinguals across all of the tasks. Since the focus of the study was different script bilinguals, the next section reviews the role of orthography in lexical access, selection, and retrieval.

2.5 The role of orthography in word production

This section discusses the role of orthography in word production, and how it affects the process, even if the written form is not present in a task.

There is a general assumption that the verbal production of a word involves the activation, or retrieval, of its syntactic, semantic, and phonological representations, but not its orthographic information (e.g., Chen et al. 2002; Roelofs 2006). This is primarily due to the written form being absent from the spoken word production process. However, several behavioural studies (e.g., Frauenfelder et al. 1990; Chéreau et al. 2007; Han and Choi 2016) concluded that the orthographic representation is active, to some degree, in all adult literate language comprehension and production activities. For instance, Frauenfelder (1990) noted the activation of orthographic representation during spoken word recognition in phoneme monitoring tasks and a similar finding was observed by Chereau et al. (2007) and Taft et al. (2008) in primed auditory lexical decision tasks. This finding was replicated across different languages with an alphabetic orthographic system, such as English (Miller and Swick 2003), French (Pattamadilok et al. 2007), Portuguese (Ventura et al. 2004), and across the non-alphabetic orthographic system, such as Chinese (Zou et al. 2012; Qu and Damian 2016). It was attributed to the fact that phonological representation is bi-directionally related to orthographical representation, namely access to phonological representation entails automatic parallel access to the orthographical representation.

The same pattern of results was observed in language studies that explored the impact of orthography on language production. For example, Lupker (1982) tested his participants with a picture-word interference task, in which they named pictures while ignoring visually presented distractors. In one experiment, the pictures and distractor words were either orthographically related but phonologically unrelated or orthographically unrelated. In the orthographically related condition, the name of the picture and the distractor word had identical spellings, except for the first letter, and their single vowel sounds were different (e.g., *foot-boot*), whereas in the orthographically unrelated condition, the words had different phonemes/graphemes (e.g., *foot-bar*). The study found that picture naming was facilitated significantly when the distractor word and the picture name overlapped in their orthography. Meanwhile, in the second experiment, three conditions were created in which the distractor words and the picture name were: (i) phonologically and orthographically related, such as *broom-room*; (ii) phonologically related but orthographically unrelated, such as *flower-hour*; and (ii) unrelated distractor words, such as *broom-truce*. The results confirmed a facilitation effect of 55 ms in the phonological and orthographical overlap condition, while there was a facilitation effect of 23 ms in the phonological overlap only condition, relative to the unrelated condition. In addition, Weekes et al. (2002) tested Chinese speakers in a picture word interference task, and found that they produced facilitation effects relevant to the unrelated condition when the distractors were phonologically and orthographically related. Thus, the findings suggest that both phonological and orthographical similarity contributed to the facilitation effects observed in naming pictures. These results were therefore further evidence of bi-directional connectivity between phonological and orthographical representation during word production.

The facilitation effect reported, due to orthographic overlap, between distractor words and target picture names, motivated the present study to argue that orthographic overlap may contribute to the cognate facilitation effect found in naming cognate pictures. Thus, this study investigated the cognate facilitation effect in different script bilinguals (i.e., with zero orthographical overlap) to determine whether orthography plays a role in the cognate facilitation effects observed in word production studies. If an orthographic overlap is essential for obtaining the assumed facilitation effect, then bilinguals whose languages possess different scripts should not reveal this effect (this was addressed in experiment one, chapter four). Such a finding would indicate that the lack of competition reported during lexical selection was not due to bilinguals attending selectively to the target language, but due to script similarity (i.e., orthographic overlap) that facilitates lexical access and the retrieval of a word in same script bilinguals.

In order to exclude any potential influence of visual word recognition, this study employed the masked priming technique, in which the prime words were fully masked and unconsciously available to the participants. The next section discusses how the different scripts might act as a language cue and modulate lexical access and selection for different script bilinguals.

2.6 The Language Cue Hypothesis

The concept of a language cue was represented in several ways in previous word production studies. For example, in Green's (1986) Inhibitory Control Model, the language cue is a tag that forms part of the lexical item and helps to identify which language it belongs to, namely either the L1 or the L2, while other researchers have identified orthography, script, and task processes, such as the translation Stroop task (e.g., Miller and Kroll 2002; Hoshino and Kroll 2008) as cues. Previous researchers argued that these cues might modulate the parallel activation of the lexical nodes in the target and non-target language, and direct lexical access to the target language (e.g., Meuter and Tan 2003). Therefore, it is a possible that the non-selective lexical access observed in same script studies is due to script similarities of the bilinguals' two languages. The present study aimed to test this view. If script differences modulate lexical access and selection, we anticipated finding no cross-language activation in our experiments, as the participants were Arabic-English bilinguals whose languages have different scripts. Next, we discuss briefly the studies that investigated the effect of language cues in production.

In their study, Miller and Kroll (2002) tested whether the same pattern of lexical competition reported in picture-word interference tasks could be replicated in a different task when a language cue is available visually to the participants. They asked English-Spanish bilinguals to perform a translation Stroop task, in which they were required to translate words from one language into another as rapidly and accurately as possible, whilst ignoring visually presented distractor words. The distractor was either semantically-related, form-related, or unrelated. Moreover, the distractor was presented either in the language of the word to be translated (i.e., in the L1 when the translation was from the L1 to the L2), or in the language of the production (i.e., in the L2 when the translation was from the L1 to the L2). The results demonstrated a semantic interference effect, along with a form facilitation effect when the distractor word appeared in the language of the production. However, no semantic interference or form facilitation effect was observed when the distractor word appeared in the language of the word to be translated. This was attributed to the presence of a language cue that eliminated lexical competition in the word translation task. Miller and Kroll (2002) argued that unlike in a picture naming task, in terms of translation, a cue was present in the target word for each language group, which thus acted to reduce cross-language activation. In order to illustrate this, the participants were instructed to translate a Spanish target word into English, and hence were aware that they should not use Spanish. This thus enabled them to ignore a Spanish distractor word easily. Hence, the study provided evidence for the language cue hypothesis in a translation Stroop task by same script bilinguals.

The current study examined these findings using a prominent language cue, namely the different script. In the priming paradigm, the prime words were masked, which should not have been an issue, as discussed previously when reviewing how orthography affects word production, even when it is not present. The study investigated whether different scripts modulate not only lexical selection, but also lexical access and retrieval. This is because previous investigations of the role of different scripts in word production (e.g., Hoshino 2006; Hoshino and Kroll 2008) reported that when the script is perceptually present, bilinguals exploit the script differences to obtain cues regarding the intended language, thus reducing cross-language activation, and engendering a more rapid retrieval of a target lexical node during single word production.

In summary, the current study expanded on known findings to investigate whether the language cue hypothesis holds true for languages with different scripts, such as Arabic-English bilinguals. It employed the masked priming paradigm in picture naming to avoid the methodological flaws associated with the other methodologies discussed in section 2.3. An in-depth review of this method is presented later in this chapter, but first, we briefly discuss the linguistic features of the Arabic language to illustrate its differences from the English language.

2.7 Linguistic Features of the Arabic Language

Arabic is a Semitic language. The term ‘Arabic’ refers to three forms of the language: (i) classical Arabic, which is the language of the Holy Quran and religion; (ii) modern standard Arabic, which is the language of education and formal spoken/written communication; and (iii) colloquial Arabic, which is the language of everyday informal oral communication (Boudelaa and Marslen-Wilson 2010). Phonetically, modern standard Arabic has six vowels (short /a/, /i/, /u/ and long vowels /aa/, /ii/, /uu/), two diphthongs (/ae/ and /ao/), and 28 consonants (Alotaibi and Meftah 2013). The duration of vowel sounds is phonemic, which is one of the major distinctions of Arabic compared to English. Each short Arabic vowel is phonetically identical to its long counterpart, which is to say the only difference is the duration, whereas in English there are more than 13 different vowels, depending on whether it is the American or British dialect, and these include several diphthongs. In terms of the Arabic text, it is written/read from right to left, and uses a cursive script, compared to English, which is written/read from left to right uses the Latin script (see appendix J for an illustration of Arabic letters). Finally, there is no distinction in the Arabic language between upper and lower cases as there is in English language. These differences indicate that the Arabic and English languages are distinct from one another, and that they share no similar features that would automatically trigger the activation of the non-target language. The prime words were presented visually in modern standard Arabic, i.e., the standard written form (in experiments one, two, and three [chapters four, five, and six, respectively]). The next section critically reviews the masked priming paradigm, highlighting how it was employed to address the research questions in the present study.

2.8 The masked priming paradigm

The present research employed the masked priming procedures in a picture naming task to test the reported effects found in the picture word interference tasks, namely the cognate facilitation effect, the identity facilitation effect, and the semantic interference effect reviewed in section 2.3. This section provides a critical review of the priming procedures, and how they are used to test these effects, including a justification for adopting this method as the main approach, for addressing the research questions of this study.

In the masked priming paradigm, a prime word is generally presented visually, and is briefly preceded and followed by a mask, which is typically a series of # symbols, with the target word subsequently presented either for a specific period, or until the participants provide a response.

Previous studies used the priming paradigm varied the status of the prime word, with some employing the visual mask (#######) to ensure the unconscious processing of the prime words, while others used unmasked prime words, and depended on the rapid duration of the prime to ensure unconscious processing. The parameters associated with this paradigm are:

(i) The modality of the presentation of the prime, either visually or auditorily;

(ii) The time course of the event (the SOA and the duration of the presentation of the prime);

(iii) The relationship between the prime and the target.

It is reported that any difference in these parameters causes a variation of performance within the priming task; for example, longer presentation times enable the participants to notice the relationship between the prime and target, and thus to use predictive strategies (Ferrand et al. 1994; Alario et al. 2000, p. 743). The most common dependent measure reported in these studies was mean response latencies (i.e., measuring participants' reaction times in ms). The next section reviews the studies that employed this paradigm, and discusses how it is suitable for investigating the semantic interference effect.

The masked priming paradigm is mainly used in word recognition studies in which a prime word precedes a target word that the participants are required to read aloud, or make a lexical decision about. In word production studies, this paradigm has typically been employed to investigate the production process by monolingual speakers. The investigations mainly manipulated the semantic relatedness between the prime word and the target picture, in order to investigate the lexical selection process. For example, Alario (2002) asked French speakers to name a picture that was preceded with a prime word that remained on screen for 100 ms, followed by a blank screen for 14 ms, (a short instance SOA, prior to the picture being presented). An inhibition effect was observed when the picture and the prime word were semantically related. Similarly, Bajo et al. (2003) manipulated variables so that the prime words were masked with a forward and backward mask, and were either semantically related to the name, or unrelated. The prime duration was also manipulated at 50, 75, or 100 ms. According to Alario et al. (2000), semantic interference appears at 100 ms exposure to the prime, and thus Bajo et al. (2003) manipulated the prime duration to cover the minimum time for the effect to appear. The results confirmed a semantic interference effect at the 100 ms, but not at 75 or 50 ms, suggesting a critical processing time. In other words, primes presented for 50 and 75 ms are processed at the semantic level, but do not cause interference. The interference effect found at 100 ms was attributed to the co-activation of lexical nodes at the lexical level that impeded the selection process. The current study employed masked priming, and challenged the semantic-interference effect by examining the performance of bilinguals using different scripts. We now extend our discussion to explain why masked priming is suitable for investigating the identity effect.

As discussed in section 2.3.2, the identity effect (i.e., reported in picture-word interference task) is located at the lexical level, and occurs due to the activation of the non-target lexical nodes (i.e., it is triggered by the translation distractor word), which in turn sends extra activation to the target lexical node, facilitating its selection. The present study argued that the identity effect can be tested in the masked priming paradigm, as its underlying process is like the picture word interference task, with the exception of the visual availability of the prime/distractor word. Several previous word recognition studies employed the masked priming paradigm to investigate the within language identity effect in monolinguals (i.e., the same word is presented as a prime and as a target word) and the results showed that the masked prime word activates its corresponding lexical representation at the lexical level, which causes faster lexical decision processing (e.g., Forster and Davis 1984).

Moreover, in a pioneering monolingual word production study, Ferrand et al. (1994) combined the masked priming and the picture naming task, and found that an identical prime facilitates the naming of the target picture. Inspired by Ferrand et al. (1994), Chauncey et al. (2009) replicated the study by testing English-French bilinguals using ERP recordings (defined in chapter two, section 2.3.5). The participants were asked to name a picture of a common object in their L1 or L2, and the picture was preceded by a brief visual presentation of a masked prime word, half of which were in the participants' L1 (i.e., English), and the other half in their L2 (i.e., French). The prime words' language varied from trial to trial within a block, namely switch trials, and consisted of: (i) the name of the picture (i.e., in English or French) in the related condition; for instance, the words '*milk'* and '*lait*' primed a picture of a gallon of milk; or (ii) the name of a different object, in English or French, in the unrelated condition; for instance, the words '*drum'* and '*tambour'* preceded a picture of a gallon of milk. The prime words were presented for a duration of 70 ms. The results demonstrated a significant facilitation effect resulting from the same name primes, namely repetition primes, as well as the translation primes. This finding supported the argument that, despite being masked and presented only briefly, the prime word activated its corresponding semantic and lexical representation, thus increasing the speed with which the picture names were retrieved. So far, this section has discussed how the priming paradigm is suitable for investigating the semantic interference and the identity effect. Next, we discuss how this paradigm can also test the cognate facilitation effect.

The cognate facilitation effect found in picture naming task is attributed to the phonological overlap between the L1 and L2 picture name, as discussed in chapter two (section 2.3.1). In the picture naming task, there is no visual presence of the target word or the non-target word, yet facilitation was observed, and interpreted as evidence for the language-specific view and the cascaded manner of activation. Using the masked paradigm, the present study aimed to raise the activation level of the phonological representations of the non-target lexical node by priming the picture with its L1 name. It was intended that this would enable the current views of lexical selection and the flow of activation to be tested, as detailed later in chapter four (experiment one). The question of interest was whether a masked cognate name would activate its phonological representation at the phonological level? According to previous word recognition studies, the recognition of a target word, such as *made*, is facilitated when preceded by a briefly presented phonologically related prime word, such as *maid*, or a non-word, such as *mayd* (Dufour 2008). This therefore suggests that the prime word activates its phonological representations. Moreover, previous word production studies argued that the "phonological priming effect may shed light on the nature of bilingual lexicon and the lexical retrieval in speech production" (Collins and Ellis 1992, p. 376). In their work, Collins and Ellis (1992) requested that their English speaking participants repeated aloud a number of auditorily presented prime words, followed by named picture targets. The relationship between the prime words and the name of the pictures was manipulated, with the prime and the name of the target picture either phonologically related or unrelated. They also tested whether the position of the shared phonemes had the ability to modulate the phonological priming effect. More rapid naming latencies were observed for target words that shared phonemes in the same position, relative to the unrelated target words. The identification of a phonological facilitation effect suggested that the prime word activated its corresponding representations at the lexical level and at the phonological level. Thus, it was plausible to employ it in our investigation of the process of cognate naming, as this process involves not only lexical activation, but also phonological activation. Therefore, the present study used the masked priming paradigm as a new technique to investigate the manner of lexical selection when naming cognates and the flow of activation in different script bilinguals.

In conclusion, masked priming methods were primarily employed in previous studies to investigate visual word recognition by monolinguals and bilinguals, with, to the best of the current researcher's knowledge, only a few studies using this paradigm to investigate word production in monolinguals (e.g., Alario et al. 2000; Bajo et al. 2003; Finkbeiner and Caramazza 2006), and one study in same script bilinguals (Chauncey et al. 2009), although they did not address the problem of lexical access, selection, and retrieval. Thus, the present study employed masked priming procedures to access the different stages of the word production system in bilinguals, and to identify whether the non-target lexical nodes are phonologically encoded. In addition, this study employed the animacy decision task, the inclusion of which is discussed briefly in the next section.

2.9 The animacy decision task

The animacy decision task involves conceptual processing in which the participants indicate via a button press whether a target picture represents a living or non-living object. According to Bugaiska et al. (2019, p. 882) "animates are living things that are capable of independent movement and can suddenly change direction without warning", such as animals and humans (i.e., they are self-propelled, unlike robot and vehicles). The present study adopted this definition and the participants were informed that they should respond to the task according to this criterion. Animacy decision is a conceptual task that requires deep semantic processing (Fliessbach et al. 2010), as the participants must focus on the meaning of the stimuli during the task. It was employed previously in several semantic memory studies investigating the context-dependent nature of semantic memory (Pecher and Raaijmakers 2004), and the effects of semantic processing on enhancing memory encoding (Fliessbach et al. 2010). The common finding of previous animacy decision tasks is that animate objects are recognized faster than inanimate objects, namely the animacy effect (New et al. 2007; Pratt et al. 2010), and it was argued that this is because animate objects, such as rabbits, sheep, and persons, are more important for fitness and reproduction than inanimate objects (e.g., mountain, fork) (Gelin et al. 2017).

The animacy decision task was employed in the present study to help to identify the locus of the semantic interference effect (experiment three, chapter six). In this study, the animacy decision task enabled access to the conceptual processing only, which was considered a possible candidate for the locus of the effect. To the best of the author's knowledge, this task has never been used to investigate the locus of semantic interference effect in bilingual's word production.

The next chapter discusses the research methods used in the five experiments included in this study, along with the research questions addressed.

Chapter 3 An Overview of the Experiments and General Methods

As detailed in chapter one, the main objective of this study is to generate an understanding of how lexical access is achieved by adult bilinguals (Arabic/English) using different scripts and to advance a theory that accounts for contradictory findings in the literature regarding the manner of language selection (i.e., language specific vs non-specific selection), and the flow of activation (i.e., discrete vs cascaded flow). To this end, we re-investigated the commonly reported effects on lexical access i.e., the semantic interference\facilitation effect, cognate effect, identity effect, and the phonological interference effect. Five experiments were conducted implementing three different methods: masked priming in a picture naming task (experiments one, two and three), the animacy decision task (experiment three), and the phoneme monitoring task (experiments four and five) which were evaluated (see section 2.8, 2.9, and 2.3 respectively) as being the most robust approach for the investigation. These experiments were conducted in a sequential order. The findings of each experiment influenced the design and objectives of the next experiment until the researcher was confident that the research questions had been addressed fully.

A critical evaluation of the literature determined that the masked priming task in picture naming was the best methodology to tap into the cognate effect, the semantic interference effect, and the identity effect. The animacy decision task was employed to locate the main cause of the semantic interference effect and, the phoneme monitoring task was adopted to further examine whether the activation of non-target lexical nodes cascaded to the phonological level or not. The five experiments addressed different part of the research questions as shown in Table 1, which summarises the experimental conditions and main objectives of the experiments which are detailed further in chapters four, five, six, seven, and eight.

Exp	Task	Prime Type	Participants	Goals
Exp1 (Ch. 4)	Naming cognate and non-cognates pictures in a masked priming task	• L1 name of the picture • Unrelated prime word	Two groups of Arabic-English bilinguals: Highly proficient and less proficient	Investigate the following: • lexical access in different script bilinguals. • The manner of lexical/phonological selection • The flow of activation (cascaded vs discrete). • The role of proficiency level in cross language activation and selection.
Exp 2 (Ch. 5)	Naming non-cognate pictures in a masked priming task	• Semantically related words • Unrelated words	Two groups of Arabic-English bilinguals: Highly proficient and less proficient	Investigate the following: • lexical access in different script bilinguals. • The manner of lexical selection. • The role of proficiency level in cross language activation and selection. • Whether the semantic interference effect can be found in masked priming picture naming task.
Exp 3 (Ch.6)	Naming non-cognate pictures in a masked priming task vs animacy decision task	• Semantically related words • Unrelated words	Two groups of Arabic-English bilinguals: Highly proficient and less proficient	Investigate the following: • The locus of semantic interference effect. • Lexical access in different script bilinguals. • The manner of lexical selection. • The role of proficiency level in cross language activation and selection.
Exp 4 (Ch.7)	Phoneme monitoring task: using shared phonemes across the two languages.	• Phoneme in the L1 name of the picture. • Phoneme in the L2 name of the picture • Phoneme that is not part of the L1 or L2 name of the picture.	Two groups of Arabic-English bilinguals: Highly proficient and less proficient.	To investigate the following: • Lexical access in different script bilinguals. • Flow of activation (cascaded vs discrete). • The role of proficiency level in cross language activation and selection.
Exp 5 (Ch. 8)	Phoneme monitoring task: using distinct and shared phonemes across the two languages	• Phoneme (shared) in the L1 name of the picture. • Phoneme (distinct) in the L1 name of the picture. • Phoneme in the L2 name of the picture. • Phoneme that is not part of the L1 or L2 name of the picture.	Two groups of Arabic-English bilinguals: Highly proficient and less proficient.	To investigate the following: • Lexical access in different script bilinguals. • Flow of activation (cascaded vs discrete). • The role of proficiency level in cross language activation and selection.

Table. 1 Summary of the five experiments.

3.1 Study participants

This section provides a general description of the participants recruited across all experiments. A total of 466 Saudi females[10] who speak English as their second language were recruited. They have normal or corrected-to-normal vision and reported no learning difficulties, such as dyslexia or dyspraxia. Their ages ranged between 19 and 43 years. They were all bilingual speakers as those who speak more than two languages were excluded. It is important to note that, each participant took part only once in this study so that there was no overlap of participants across the experiments. The participants were English language teachers or students recruited from two different university language institutes in Saudi Arabia. The language institutes are restricted to the education of preparatory year students, i.e., first year college students who are required (in addition to other basic courses in several different subjects) to follow comprehensive English language courses, to enhance their English language skills and facilitate their college entry. Following the successful completion of this preparatory year, the students select their major and enter their chosen department. Enrolment on these English courses was based on the students' IELTS scores, namely students with overall average scores of 2.5 to 5 were required to take these comprehensive English language courses, whereas those who had a score of greater than 5 were exempt. Those enrolled on English language courses formed the less proficient group in this study.

On the other hand, the English language teachers were holders of MAs in TESOL, TEFL or Applied linguistics. The majority had taken the CELTA (Certificate in Teaching English to Speakers of Other Languages) examination from the Cambridge English Language Assessment, which forms part of the Cambridge University. Furthermore, 90% of these teachers had taken part in various language teaching courses provided by the Cambridge

[10] The participants recruited in the five experiments were all females, because we had permission to access only the female campus to collect data for this study in Saudi Arabia.

Institute. Their IELTS scores ranged from between 6 and 9, and they comprised the highly proficient group.

In every experiment, a different group of less and highly proficient participants was recruited, and their IELTS scores were tested statistically to ensure that the two groups differed significantly with regard to their proficiency level (see chapters four, five, six, seven, and eight). In addition, further proficiency measures were administered, including a language history questionnaire involving a self-assessed proficiency rating and a lexical decision task, as detailed in the next section. Following a description of these measures, the ethics procedures for this research are detailed in section 3.2.

3.1.1.1 Proficiency measures

3.1.1.1.1 IELTS Test

All the English language learners had taken the International English Language Testing System (IELTS). The English language teachers had also taken the IELTS test, with approximately 80% possessing recent test scores, so that they could apply to undertake PhD programmes in the UK, while the remainder had a score, gained two or three years previously, due to having recently acquired their MA degree. In this research, the IELTS score formed the primary measure of proficiency alongside a language history questionnaire which was given to provide information about the linguistic background of the participants.

3.1.1.1.2 Language history questionnaire

A language history questionnaire was distributed to all of the participants, in order to collect information about their L1 and L2, namely their daily usage of their L1 and L2, the amount and type of their L2 language learning experience, their age of L2 acquisition, and how long they had lived in an English-speaking country. It also asked whether or not the participants spoke more than two languages, and if this was the case, they were instructed not to complete the questionnaire. Lastly, the questionnaire asked the participants to rate their L2 proficiency level under the four language skills: reading, writing, speaking, and listening. A blank copy of the language history questionnaire is provided in appendix C. The next section discusses the lexical decision task employed in this study (experiments three, four and five), including the task procedures, the selection of material used, and the analysis of the data collected.

3.1.1.1.3 Lexical decision task

Most bilingual studies, within experimental psychology (e.g., Jared and Kroll 2001; Prebianca 2014) rely on participants' self-ratings of proficiency and language background questionnaires as the only source of proficiency information which we did in experiment one and two, along with their IELTS scores. As discussed in section 2.4, it is recommended that combined proficiency measures are adopted to increase the reliability of the classification of the participants' proficiency level.

In experiments three, four, and five, a lexical decision task was used as an objective measure of proficiency level, because these experiments were executed online, due to Covid 19 restrictions, and not in labs as experiments one and two (see section 3.4) were, and thus it was necessary to employ this additional measure of ability as the researcher did not meet the participants in person. One of the advantages of this online procedure was that the participants were more available than they might be otherwise, and it was possible to set appointments at more convenient times for both the participants and the researcher than would be the case with the lab-based experiments, for which the timings were subject to lab availability and the participants' study/teaching schedule.

The lexical decision task is a standard psycholinguistic tool for measuring word recognition skill and is considered an effective measure of vocabulary size in L2. It has been used in several studies as a measurement of proficiency level (e.g., Hermans et al. 1998; Huibregtse et al. 2002). It is a forced-choice categorisation task in which an individual is presented with a set of lexical items consisting of words and non-words. The non-words used in the test are pseudowords, which are phonologically and orthographically possible words in the target language. For each item, participants decide whether the item meets some pre-specified criteria, usually whether it is a word in the language or not (Harley 2001). Accuracy (Acc) is measured by the ability to distinguish between the two item types, and speed is measured by the time it takes to respond to each item i.e., reaction times (RTs). A high rate of accuracy indicates a large vocabulary size, and a faster response time indicates the development of L2 processing skills (Segalowitz and Hulstijn 2005; Mochida and Harrington 2006). Hence, the highly proficient group was expected to provide more accurate responses and faster performances than the less proficient group.

a. Material selection for the lexical decision task

The material was sampled from a list used by Azuma and Van Orden (1997) in their lexical decision task. It consisted of forty-eight English words and forty-eight English pseudo-homophones. The words varied in number of meanings (i.e., few or many). In Azuma and Van Orden's norming study, a word with four or less meanings was classified as few, while a word with six or more meanings was classified as many. Also, the words varied in the relatedness of word meanings (low, high) i.e., words with relatedness scores < 3.0 were classified as low and, words with relatedness scores > 3.5 were-classified as high. So, there were four categories of words that were included in our task: twelve words with few meanings and low relatedness (e.g., bark), twelve words with few meanings and high relatedness (e.g., rake), twelve words with many meanings and low relatedness (e.g., plot), and twelve words with many meanings and high relatedness (see full list in appendix D). All non-words were pseudo-homophones i.e., non-words that share pronunciation with real words (e.g., treet).

To ensure that no more than three words from the same condition (e.g., many-low) were presented in a row, we created four stimulus lists in which each list had 12 words and 12 non-words. The 12 words consisted of: three (many-low), three (few- high), three (many- high), and three (few-low). The lists were counterbalanced across participants (full list can be found in appendix D).

b. Lexical decision task procedure

The task was performed online due to the outbreak of the Covid-19 pandemic which meant that physical labs were shut and potential participants could not leave home. The task was created using PychoPy 3 which is a free cross-platform package used to run experiments in behavioural sciences (further description is provided in section 3.4.2). Responses were collected via mouse click or touch screen clicks. Each participant attended a video conference meeting with the researcher in which she received oral instructions in English about the task. We did not use the participants' L1, in order to avoid activating the L1 lexicon before commencing the task. Our objective was to observe the effect of L1 prime words on the L2 production. Thus, it was necessary not to establish any conversation in the L1 ahead of the tasks. This was applied to all tasks' procedures employed in experiments one, two, three, four and five. Following ethical clearance (see section 3.2), and after receiving instructions, a training and task link were sent to participants. They used their own computers/tablets to perform the task. They were informed that strings of letters would be presented one at a time on the screen and they had to decide whether each string of letters was a real English word or not. A training link with 6 trials was sent first to familiarize participants with the task. Once the researcher received confirmation from participants that they were ready to start, the link to the main task was sent. To ensure no disturbance would occur during the task, participants were instructed to be in a quiet room, to stay focused as much as possible, and to turn off their camera and microphone during the performance of both the training and main task. The researcher's camera and microphone were turned off as well. On each trial, a fixation sign (+) was presented for 500 ms at the centre of the screen. When the fixation sign was replaced with a string of letters, participants were required to judge whether the string of letters was a real English word or not. If it was an English word (e.g., hide, lock, etc.), they clicked (mouse click or screen touch click) the "yes" button on screen; if it was not a real English word (e.g., korn, frum, etc.) they clicked the "no" button on screen. After they responded, a fixation sign appeared for 500 ms again and another string of letters was presented.

c. Data analysis of the lexical decision task

Reaction times (RTs) were calculated only for correct responses. Reaction times that were 2.5 standard deviations above or below the mean were identified as outliers (i.e., extreme values) separately for words and nonwords and excluded from the analyses. This was done as it is recommended to remove outliers, in order to maintain the integrity of the mean, as the statistical tests used in this study are sensitive to outliers. Then, the final mean reaction times and accuracy for the correct responses were calculated for each condition, and these results enabled the researcher to place the participants in the high or low proficiency group. In the next section, we discuss the research ethics of this study, and how the study's participants were approached and recruited.

3.2 Research ethics

The ethical approval for this study was obtained from the Research Ethics Committee at the School of English, Communication, and Philosophy (ENCAP) at Cardiff University, then consent was obtained from the relevant institutions in Saudi Arabia, as well as every participant individually. In order to obtain ethical approval to conduct the study, the researcher first attended a mandatory training course on Research Integrity, provided by Cardiff University, which seeks to help researchers understand their responsibilities, and ensure that their research is conducted to the highest professional standards. The researcher then prepared and submitted to the ENCAP Research Ethics Committee the following: (i) an application form that included a detailed description of the intended study (i.e., the research objectives and questions; a description of the targeted participants and the type of information to be obtained; and the data collection method, including the procedures, and the handling of anonymous data); (ii) a participant information sheet that explained the purpose of the research and what the participants would be required to do, how they would be involved, and how their data would be confidentially stored; (iii) the recruitment letters for the five experiments and the online survey (the semantic and phonological similarity rating), which introduced the participants to matters including the general objectives of the study, and the participants' role and benefits/risks, if applicable; (iv) the consent form that the participants were required to read carefully and sign; and (v) a list of written stimuli words, and a sample of picture stimuli. Once ethical approval from the Research Ethics Committee at Cardiff University had been granted, the researcher contacted the language institutes at the two universities in Jeddah, Saudi Arabia, to seek approval to recruit the participants and conduct the data collection. The approval letter from the Research Ethics Committee at Cardiff University, and the fact that the researcher was a staff member at the two universities in Jeddah was sufficient to grant approval. Details concerning how the participants were approached and recruited is provided in the next section.

The data collection method was lab-based initially (i.e., experiments one and two [chapters four and five, respectively]), but due to the outbreak of the Covid-19 pandemic, online methods were adopted subsequently to conduct the last three experiments (i.e., experiments three, four, and five [chapters six, seven, and eight, respectively]). The change to the process was reviewed by the Research Ethics Committee at Cardiff University, and continued ethical approval was granted. For the lab-based experiments, the participants were recruited via an email that provided an overview of the experiments. The email was sent to the Head of Scientific Research Unit at the university, who oversaw the circulation of the recruitment emails to the relevant staff and students at the institutes. Upon receiving a positive response, appointments were then set for the individuals willing to participate in the study. A follow-up email that confirmed the appointment and its location (lab-room) was sent to the participants individually. The researcher orally presented the participants with a brief overview of the experiment, in order to guarantee that all the participants received the same information. They were then given a participant information sheet, which detailed how their data would be anonymously used/stored. They were required to read it carefully and ask for clarification, if needed. If they were still willing to participate in the study, they were required to read and sign the consent form. For the online-based experiments, the participants were approached and recruited via email, as for the lab-based experiments. Those who were willing to participate received an email individually, with the participant information sheet (pdf file) and the consent letter (see appendix A and B for an example of these). The email emphasized the fact that the information sheet, along with the consent form, must be read carefully before giving consent to participate in the study. If the individuals approached had any further questions, they were asked to contact the researcher via email. Once, the electronic consent form was signed and returned, an individual Blackboard video conference call was set up between the participant and the researcher, in order to perform the experiments. The participants were informed that participation was entirely voluntary, and that they were free to withdraw their consent to participate at any time, prior to the data analysis phase, which commenced seven days after their completion of all the tasks required, without giving a reason, even after signing the consent form. All the information collected from (or about) the participants during the research project was kept confidential (i.e., it could not be traced back to the individual participants, and the data was accessible only by the researcher), and any personal information they provided was managed in accordance with data protection legislation. Cardiff University was the Data Controller, and as a university is committed to respecting and protecting their personal data, in accordance with their expectations and with data protection legislation. The next section provides further details on how data was stored and managed.

All the information collected by this study was stored in such a way that it could not be traced back to participants. With the exception of the signed consent forms, the researcher anonymized all the participants' data (age, gender, nationality, education and linguistic background, button-press responses, and reaction times) collected for the research project, from the point at which the participants conducted the tasks involved. In addition, the participants' vocal responses were held confidentially and were accessible only by the researcher, and were retained for a period of three months after commencing the tasks, then were permanently deleted. All anonymized, confidential data was stored under each participant's ID number in the researcher's space on Microsoft OneDrive (a secured, password protected drive provided by Cardiff University). A scoring sheet was used to represent the data collected from the vocal recordings, and these records were processed anonymously for interpretation in the data analysis phase, later stages of this research project, and in future journal publications and presentations. The anonymised information and the signed consent forms will be retained for a minimum of five years and then will be permanently deleted (i.e., the paper records will be shredded and the digital records will be permanently deleted). The next section elaborates on how data were removed upon participants' withdrawal from this research project.

If any of the participants decided to withdraw from this research project (during the seven-day grace period for withdrawal), they were simply required to provide the researcher with their participant number. If they performed the tasks online, the participant number was the ID number they created and entered at the beginning of each task. If they performed the tasks in the lab, it was the ID number that the researcher provided them with, which referred to their order of participation. The researcher did not retain any record that connected this number to the participants. Thus, the participants were required to remember this number to give the researcher should they wish to withdraw their contribution. As stated earlier, withdrawing was possible at any point prior to the data analysis phase, which commenced seven days after the participants' completion of all the tasks required. At that point, the researcher discarded the participants' numbers from all scripts, which meant that the researcher could not identify and retrieve the individual files.

The next section discusses the two main paradigms employed in this study to address the research questions, namely the masked priming paradigm and the phoneme monitoring task. As the animacy decision task was used once in experiment three (chapter six) for a short role, it was considered to be a secondary paradigm, and it is thus discussed in the related section (3.3.1.3). A brief description of these tasks is provided, along with the objectives and the related hypothesis; more detail is given in the chapters for the relevant experiments.

3.3 Paradigms employed in this study

3.3.1 The First paradigm: the masked priming of picture naming task (Experiments one, two, and three)

This paradigm was implemented in three different experiments (one, two, and three [chapters four, five and six]), in order to investigate lexical access/selection and the flow of activation in different script bilinguals. It also tested whether or not these processes were affected by the differences in the participants' proficiency level. The paradigm enabled the testing of the cognate effect, the identity effect, and the semantic interference effect reported in the picture word interference task (defined in sections 2.3.1, and 2.3.2). As discussed earlier, this paradigm was employed due to the presence of a mask that promoted the unconscious processing of the L1 prime words and allowed us to investigate the production process free from any linguistic influence (see section 2.8). Two different types of pictures were used i.e., cognate and non-cognate pictures. The prime-target relationship and the presentation time of the prime words were manipulated based on the objective of each experiment. The highly and less proficient Arabic-English bilinguals were asked to name the pictures in their L2, and were unaware of the existence of the masked priming words. The next section summarizes the three experiments, which are detailed in chapters four, five, and six, and discusses how the masked prime paradigm was implemented in the investigations.

3.3.1.1 Experiment one: (Naming cognate and non-cognate pictures)

Experiment one investigated the commonly reported cognate facilitation effect (see section 2.3.1) in picture naming tasks, and the identity effect (see section 2.3.2) in picture word interference tasks. As discussed in section 2.3.1, the cognate facilitation effect typically supports the view that the manner of lexical selection is language specific, and that the flow of the activation of the lexical nodes in both languages cascades to the phonological level. The identity effect was typically reported in previous picture-word interference tasks, and the finding of a facilitation effect challenged the view of lexical competition, due to semantic similarities.

However, since these effects were typically reported in studies that used the picture-word interference task, and tested same script bilinguals, further examination of the underlying process was required. Specifically, as reviewed in chapter two, this study questioned the findings of the picture word interference task, due to its design flaws, i.e., the presence of the visually presented distractor word might confuse the participants and cause them to process it for production, prior to the processing of the target picture, hence the interference effect (Finkbeiner and Caramazza 2006). Thus, experiment one of the present study adopted a different methodology, namely the masked paradigm, which similarly requires a picture to be named, and a non-target word is introduced. However, unlike the picture word interference task, the non-target word is unavailable for conscious processing as it is masked, ensuring that the participants prepare vocal responses to the target picture only. The presentation time of the masked prime word for the investigation of the cognate effect and identity effect was set at 50 ms, because previous monolingual studies reported a priming effect for repetition and phonological priming at short presentation times (e.g., Ferrand et al. 1994).

As discussed in section 2.3.1, the cause of the cognate effect was attributed to the shared phonological segments across the two languages (Costa, Caramazza, et al. 2000). Thus, we hypothesized that, if the flow of activation cascades to the phonological level, then priming a cognate picture with its masked cognate name in L1 (i.e., the critical condition[11]), would induce

[11] We employ the use of the term 'critical condition' to refer to the condition in which we manipulated the relationship between the prime and the target picture, whereas the term 'control condition' is used to refer to the condition in which there is no relationship between the prime and the target picture.

an interference effect or at least slow down the speed of naming. This is because the non-target lexical nodes should be highly active, and thus impede the selection process.

With regard to the identity effect, it was hypothesized that if lexical selection is language specific, then the facilitation effect should be found when a non-cognate picture is preceded by its L1 name. However, if lexical selection is language non-specific, then an interference effect should be found because of the competition between the activated lexical nodes in the target and non-target language at the lexical level. We ensured that the prime and target picture did not overlap phonologically.

The data analysis and discussion of the results is presented in chapter four; the research questions are presented next.

3.3.1.1.1 Research questions for experiment one

- What is the manner of phonological selection/activation by different script bilinguals?
- What is the manner of lexical selection by different script bilinguals?
- What is the effect of different script on lexical access and manner of selection?
- What is the effect of proficiency level on lexical access and manner of selection/activation?

3.3.1.2 Experiment two: (priming non-cognate pictures with semantically related words)

As stated in chapter two, the semantic interference effect found in the picture word interference task was considered as evidence supporting the view of language non-specific selection. However, since the finding was based on investigations carried out using picture word-interference tasks, they can be challenged (Finkbeiner and Caramazza 2006), as described in section 2.3.2. Therefore, there was a need to test the semantic interference effect using the masked priming technique.

Thus, in experiment two (chapter five) we compared the performance of highly and less proficient bilinguals in a masked priming picture naming task manipulating the semantic relationship between the masked L1 prime word and the target picture. The time of the presentation of the masked prime words was also manipulated based on previously published findings (see section 2.8). Therefore, different presentation times (50 ms, 75 ms, and 100 ms) were employed to determine whether the effect would manifest at avsimilar SOA or earlier in bilinguals using different scripts. We avoided using a longer presentation time as it may allow participants to recognize the prime word; even more they might be encouraged to translate the prime word into the target language as discussed in section 2.8.

In this experiment, an L1 semantically related masked prime word preceded a target picture and participants were asked to name the picture in L2 as fast and accurately as possible. The mask and the short presentation time of the prime word would make the prime word unconsciously available for processing, thus minimising the chances of any sort of confusion. If lexical selection entails competition between target and non-target lexical nodes, longer naming latencies should be found in the critical condition. More specifically, the L1 prime word would send activation to the corresponding lexical node in the non-target language (L1) and thus increase its level of activation which should lead to competition. However, if lexical selection does not entail competition, naming latencies should not be affected in the critical condition.

The data analysis and discussion of the results of this experiment is presented in chapter five; the research questions for experiment two are listed below.

3.3.1.2.1 Research questions for experiment two

- What is the manner of lexical activation/selection by different script bilinguals?
- What is the effect of script differences on lexical access and manner of lexical selection/activation?
- What is the effect of proficiency level on lexical access and manner of selection/activation?

3.3.1.3 Experiment three: (masked priming vs animacy decision task)

This experiment was conducted to investigate our hypothesis about the discrepancy in the findings between experiment one and two regarding the manner of lexical selection. Strikingly, the semantic interference effect is the only indicator of lexical competition and thus, it is important to investigate the locus of this effect to determine whether it originates at the lexical level or at the conceptual level. To this end, we compared the performance of our Arabic-English bilinguals in both proficiency groups in two different tasks in experiment three: masked priming and an animacy decision task. The animacy decision task (reviewed in section 2.9) involved conceptual processing where participants had to indicate via button press whether a target picture represented a living or non-living object. The other task (the masked priming in picture naming task) involved conceptual and lexical processing i.e., target pictures in both tasks were primed with a semantically related masked word in L1. If the semantic interference effect originates at the conceptual level, we hypothesis longer reaction times in both tasks as they both involve conceptual processing. However, if the semantic interference effect originates at the lexical level, we hypothesis longer reaction times in the picture naming task only as this task involved lexical processing. The data analysis and discussion of the results is presented in chapter six; the research questions for experiment three were as follows:

3.3.1.3.1 Research questions for experiment three

- What is the locus of the semantic interference effect?
- What is the effect of proficiency level on lexical access and manner of lexical selection/activation?
- Do script differences modulate the cross-language activation in both tasks?

The next section discusses the second paradigm employed in this research, namely the phoneme monitoring task, and summarizes experiments four and five.

3.3.2 The second paradigm: the phoneme monitoring task (Experiments four & five)

The phoneme monitoring task (reviewed in section 2.3.3) was employed in experiments four and five to investigate lexical access and the flow of activation in different script bilinguals. It also examined whether the participants' proficiency level modulated lexical access and the flow of activation.

In the three masked priming experiments (chapters four, five and six), we manipulated the lexical properties of the prime word and we observed the outcome of this lexical activation at the lexical/phonological level. Although the priming method is a good candidate for investigating cross-language lexical activation, this study employed a different paradigm that would tap specifically into the sub-lexical processing i.e., the cross-language phonological activation, in order to examine whether the activated non-target lexical nodes would spread activation to the phonological level. Since the cognate effect (investigated in experiment one) is based on the existence of shared phonemes, it was necessary to investigate the cross-language phonological activation, in which there is no phonological overlap, before making any generalization. The phoneme monitoring task was employed to examine the cross-language activation with zero phonological overlap. The next two sections detail the manipulations administered in each experiment, and the related hypothesis.

3.3.2.1 Experiment four: (phoneme monitoring task: shared phonemes in L1 & L2)

In experiment four (detailed in chapter seven), highly and less proficient Arabic-English bilinguals were presented with target pictures that were preceded by phonemes (shared by both languages) and they had to decide whether the phoneme was part of the English picture name or not. For the picture of *a lion*, for example, the phonemes were of three types: (i) part of the L1 name of the picture in the critical condition (e.g., /s/); (ii) part of the L2 name of the picture in the positive [12] condition (e.g., /l/); and (iii) not part of the picture name in L1 or L2, i.e., unrelated in the control condition (e.g., /k/). Only the phonemes that are legal in both languages were employed. In total, 14 phonemes that are similar in Arabic and English, in terms of place and manner of articulation (cf. Alotaibi and Meftah 2013), were selected. They were consonants with a single letter-representation that corresponds to English letters (see appendix J for full list). Although the participants were expected to monitor every syllable of a picture name, we ensured that the target phoneme appeared at the onset position in the critical and positive conditions to avoid the effect of different syllabic positioning. We hypothesized that if the phonological segments of the non-target language are active, then longer reaction times should be observed in the critical condition. However, if the phonological segments of the non-target language are not activated, then reaction times in the critical condition should not be different from those in the control condition. The analysis and discussion of the results is presented in chapter seven; the research questions for experiment four are listed below:

[12] The terms 'positive condition' and 'critical condition' refer to the conditions in which the relationship between a phoneme and the target picture was manipulated, whereas the control condition refers to zero manipulation between a phoneme and target picture. The positive condition should elicit a positive response (i.e., click Yes), whereas the control and critical condition should elicit a negative response (i.e., click No).

3.3.2.1.1 Research questions for experiment four

- Is the flow of activation in different script bilinguals cascaded or discrete?
- What is the effect of language proficiency level on the cross-language phonological activation?

3.3.2.2 Experiment five: (phoneme monitoring task: L1 distinct phonemes vs shared phonemes)

In experiment four, we used only shared phonemes across the two languages to investigate the cross-language phonological co-activation and we observed the outcomes of this manipulation. In experiment five, by contrast, we tested whether the findings of experiment four could be replicated when a language specific phoneme (L1 distinct phoneme e.g., Arabic /ʕ/ ع) is used. So, in experiment five (chapter eight), the same phoneme monitoring task was adopted in four experimental conditions. For example, for a picture of *a spider,* four experimental conditions were created: (i) a bilingual critical condition in which the phoneme was shared across the two languages and part of the picture name in L1 (e.g., /n/), (ii) a monolingual critical condition in which the phoneme was language-distinct and part of the L1 name of the picture (e.g., /ʕ / ع) , (iii) a filler condition in which the phoneme was part of the picture name in L2, and lastly (iv) a control condition in which the phoneme was not part of the picture name in L2 or L1. It was hypothesized that if the phonological activation of the non-target segments is restricted to shared phonemes across the two languages, then longer reaction times would be observed only in the bilingual critical condition where shared phonemes that are part of the L1 name of the picture were presented. However, if parallel activation is applicable to both shared and non-shared phonological segments in the two languages, then reaction times in the two critical conditions should be longer than in the control condition. The results of this experiment are analysed and discussed in chapter eight; the research questions for experiment five are listed below:

3.3.2.2.1 Research questions for experiment five

- Are the L1 distinct phonemes (i.e., non-target phonological segments) activated at the phonological level? And what is the impact of this activation?
- What is the effect of proficiency level on the cross-language phonological activation?

The previous sections discussed the five experiments conducted in this study to test the production process of different script bilinguals, and described the two paradigms that were employed, namely the masked priming paradigm and the phoneme monitoring paradigm. The next section provides a detailed account of how these experiments were administered, along with the justification for their use.

3.4 Study experimental procedures

This study conducted the experiments involved via a face-to-face lab-based method, and subsequently an online web-based method, due to the Covid-19 restrictions that were imposed during the data collection, which meant that it was necessary for all research activity to be conducted online, or via other remote means, from March 2020 until the end of the restrictions.

3.4.1 Lab-based procedures (experiment one and two)

Experiments one and two (masked primed picture naming task) were performed in labs where the researcher recruited participants and explained the nature of the task in one-to-one meetings using the computer screens and experiments tools. Participants were tested on individual basis, and in isolated rooms. They were placed in front of high performing multimedia personal laptop, which had a 15" FHD screen and a resolution of 1920 x 1080. The viewing distance was approximately 60 cm. Both pictures and words were shown in black on a white background, and the picture size was 300 x 300 cm. All of the pictures were reproduced with sufficient clarity to be identified easily, i.e., only an image of the object was presented, with nothing accompanying it on the screen. For example, a picture of a bird showed only that particular bird, rather than a bird standing on the branch of a tree.

The participants' responses were recorded using an external electric microphone connected to the response device Chronos, which also measures reaction times (RT). The experiment was designed using E-Prime 3.0. which is widely used software in behavioural research.

The participants received oral instructions in English (reasons discussed in section 3.1.1.1.3) prior to commencing the test, with additional English written instructions being delivered on screen throughout the test. The participants were informed that they were required to look at the centre of the screen and name the pictures as rapidly and accurately as possible, and to say 'pass' if they did not know the name of the picture. The participants were also requested to avoid undesirable responses (such as 'ah' or 'uhm') when thinking of an answer, and to avoid laughing, as this would trigger the voice key. They were not informed of the presence of the primes. They had a trial training session, which was performed with separate stimuli and prime words from those employed in the actual test. The experimenter was present only during the trial sessions, to ensure that the participants fully understood the procedures and performed them according to the instructions given. The next section provides a description of the online procedures, detailing the software used for the implementation of the online tasks and how these tasks were administered.

3.4.2 Online procedures (experiments three, four and five)

The study initially planned to conduct all of the data collection face-to-face with the participants in labs, however we lost access to labs and study participants due to the outbreak of Covid-19 pandemic. Thus, experiments three, four, and five were administered online. A web-based data collection method was adopted, and we recruited different groups of Arabic-English participants who took part in the study remotely from their own computers. Revised ethics was achieved for the online data collection (see section 3.2 for detail). Lately, this form of experimental setup has become popular because researchers can easily access and collect a large amount of data from a wide range of locations and populations (Gosling and Mason 2015; Woods et al. 2015; Stewart et al. 2017). Moreover, different tools have been developed to facilitate the building of web-based experiments to ensure careful display timing of the stimuli and accurate recording of participants' responses within web-browsers (Kochari 2019). This has made web-based data collection suitable for many experimental paradigms typical in cognitive psychology (Semmelmann and Weigelt 2017; Kochari 2019). A discussion of the web-based tool that was employed is presented next, along with the justifications for its use.

Experiments three, four and five were thus designed on-line. At the time of online data collection, an online version of E-Prime was released i.e., (E-Prime Go) that allows researcher to send a task file to participants through emails. However, this online version was compatible only with Windows and not the other operation systems such as Android or OSX and has additional challenges including that the reaction times are measured by the participants' local personal computers rather than a reliable server, and thus results are subject to variability across participants owing to different computing powers and E-Prime Go is dependent on participants' cooperation and computer skills as they need to go through several steps including downloading the file, running and saving the trials, and uploading and sending the file to the researcher. Thus, we decided against this package and instead adopted PyschoPy3, a more practical and reliable option which has been used extensively, as summarised below and is proven to generate accurate data. This software (PsychoPy3 v3.1.5 Builder GUI) is a free cross-platform package that is used to run a wide range of experiments in behavioural sciences (e.g., neuroscience, psychology, psychophysics, linguistics, etc.) with precise spatial control and timing of stimuli (Peirce et al. 2019). Bridges et al. (2020) compared different software packages (including E-Prime) used to conduct online behavioural experiments to measure response times and performance of participants as well the precision and accuracy of visual and auditory stimulus timing and response times across these packages and concluded that PyschoPy 3 v3.1.5 achieved impressive reaction time precision of under 4 ms on all web browsers and thus was adopted here in this study. Next, we provide further details regarding how the tasks were administered, and what instructions were given to the participants.

After being created on PyschoPy, the tasks were exported to PsychoJS v3.1.5 and uploaded to Pavlovia.org using the GUI, where the tasks were run from a browser. PsychoPy can run efficiently on different web browsers like Chrome 76 (Windows), Chrome 75 (macOS), Firefox 68 (Windows), Firefox 69 (macOS), Safari 12 (macOS), but not on Edge at the time we executed this experiment, thus participants were informed not to use it, but we were confident that there were sufficient browsers available to ensure that participants were not excluded. Since all tasks required full-screen preview; we ensured this using the experiment settings in the builder view. The one major obstacle with PsychoPy is that its online version does not have the option to record reaction times to vocal responses or record participants' actual responses. Thus, in the masked primed picture naming task only (experiment three, chapter six), participants were asked to record their screens with audio and send the recording to the researcher. At the data analysis stage, Praat, a software for analysing, synthesizing and manipulating speech and other sounds, was used to measure voice onset time manually. In addition, a tone (a gentle mouse click) was added at the onset of the target picture which allowed us to identify the start point of measuring the reaction time to each trial in Praat. Then each trial sound file was analysed manually by measuring the length of the silence between the onset of the tone and participant's vocal response.

When conducting the online tasks (the masked primed picture naming, phoneme monitoring, animacy decision, and lexical decision tasks), participants were tested on an individual basis on a video conference call with the researcher. They were instructed to be seated in front of their own personal laptop, desktop, or tablet with a suitable viewing distance of approximately 60 cm if possible. They were asked to be in a quiet room and avoid any distractions and to minimize any sort of disturbance from family members while performing the tasks. They had all be cleared through our ethics procedures as described in section 3.2.

3.5 Selection of materials used in all experiments

3.5.1 Pictures

The pictures used in all the experiments were black and white line-drawings of objects (e.g., food, animals, furniture, body parts, clothing and musical instruments, etc.) collected from the International Picture Naming Project online-database (Szekely et al. 2003; Szekely et al. 2004) (see appendix K for pictures samples). We matched the picture names on variables that are often considered critical to the type of processing level involved in the tasks and the experimental conditions. More specifically, variables including familiarity and imageability, which are usually thought to influence the semantic processing level (Brysbaert et al. 2000), whereas frequency typically impacts the lexical-phonological processing stages (Levelt et al. 1999). Moreover, the age of acquisition of a semantic category is known to affect both the speed and accuracy of the individual's lexical/semantic processing (Räling et al. 2017). In consideration of this, the MRC Psycholinguistic database (Coltheart 1981) was utilized to match the target names for familiarity and imageability, and the International Picture Naming Project database[13] (Szekely et al. 2003; Szekely et al. 2004) to number of letters, frequency, number of syllables, visual complexity and age of acquisition.

3.5.2 Prime words

All the related and unrelated Arabic prime words, used in experiments one, two and three (chapters four, five and six), were carefully matched in terms of number of letters, number of syllables, and frequency. The data was taken from the published normative online databases for nouns, "The Gulf Arabic Nouns" (Khwaileh et al. 2018) and the Modern Standard Arabic lexical database "ARALEX" (Boudelaa and Marslen-Wilson 2010).

3.6 Data Trimming for all experiments

Each participant's reaction times data was exported to an Excel file for analysis, and the vocal recordings/ button press responses were analysed manually by the researcher.

For the masked primed picture naming tasks, the researcher listened to every trial response, scoring it as either a correct or incorrect answer. We applied a liberal criterion for accuracy, i.e., synonyms of a given non-cognate picture name in experiments one, two, and three (chapters four, five, and six respectively) were scored as correct answers. Responses that did not match the expected picture names, began with hesitation/ a 'pass' response, or were not recorded due to technical issues were scored as errors. For the animacy decision task and the

[13] We used the MRC Psycholinguistic database as a secondary source, because familiarity and imageability data were not available on the International Picture Naming Project database.

phoneme monitoring tasks, we matched the participants' responses (yes/no) against the answer key sheet prepared previously by the researcher.

For all the tasks, the reaction times for correct answers only were included in the analyses. When trimming the correct responses data, we employed two methods that are applied widely in psycholinguistic research; namely the absolute cut-off and the standard deviation rule. Following the absolute cut-off rule, we trimmed responses that were less than 300 ms or greater than 3000 ms, as typically applied in most picture naming studies (e.g., Hoshino 2006; Hoshino et al. 2021). Then, reaction times that were 2.5 standard deviations above or below the mean were identified as outliers, and excluded from the analysis. The percentage of errors and outliers excluded from the analyses are reported in each experiment (chapters four, five, six, seven, and eight). For each of these experiments, an ANOVA analysis was performed for the mean response latencies and accuracy per subject and item. The participants' mean reaction times for their correct responses in each condition were obtained and used for the subject analysis. The item analysis was based on the means for correct responses to each item under each condition.

The next chapter reports the first experiment in this study which is naming cognate and non-cognate pictures in a masked primed paradigm. A full account of the experimental procedures, analysis of results and discussion of the major findings are presented in chapter four.

Chapter 4 Experiment 1: Masked Priming of Cognate and Non-Cognate Pictures in a naming task

This experiment investigated the manner of lexical access by different script bilinguals, namely Arabic-English speakers, to determine whether the previous findings of non-selective lexical access in same script bilinguals (Green 1986; De Bot 1992; Poulisse and Bongaerts 1994) is also applicable to bilinguals with distinct scripts. Moreover, it examined whether the flow of activation cascaded to the phonological level in different script bilinguals. Section 2.2.3 discusses the fact that the present study argues against there being a fixed locus of selection, and instead proposes that there is a selection mechanism at every representation level. Thus, this experiment investigated the manner of lexical and phonological selection to determine if competition occurs. Finally, it investigated the participants' language proficiency level, in order to determine whether it plays a role in cross-language activation, the manner of the lexical/phonological selection and the flow of activation.

To address these questions, we used cognate and non-cognate pictures for a naming task preceded by L1 masked prime words. As discussed in section 2.3.1, in the majority of word production studies, naming cognates and non-cognates were employed in picture naming tasks (e.g., Hoshino 2006; Hoshino and Kroll 2008) and picture word interference tasks (e.g., Costa and Caramazza 1999) to assess cross language interaction in bilinguals. More specifically, they were used to investigate whether the non-target lexical node (i.e., non-selected) spreads activation to the corresponding phonological segments. The typical finding in the previous literature is that cognate pictures are named faster than non-cognates, and this was taken as evidence supporting the non-selective access and a cascaded flow of activation.

However, it is unclear yet what the implication of this co-activation of the phonological representations at the phonological level is; that is, do they facilitate or hinder the phonological selection of target segments? According to Costa, Caramazza, et al. (2000), similar phonological segments in the non-target cognate name facilitate the selection of the target segments, but what if the non-target phonological segments were to receive extra activation and become highly active, will they interfere with the selection of the target segments? In other words, is the phonological selection sensitive only to the activation level of the target segments, and does it not consider the activation level of the non-target segments and will this cascaded activation manifest if the two languages shared different scripts?

To answer these questions, we explored Costa, Caramazza, et al.'s (2000) theory with regard to the manner of cross-language phonological activation and selection (detailed in section 2.3.1). In experiment one, the cognate picture was primed with the non-target name (i.e., the name of the picture in the non-target language) to increase the activation level of the non-target lexical nodes at the lexical level. If activation of non-target lexical nodes flow from the lexical level to the phonological level as assumed, the phonological segments corresponding to the non-target word should also be highly activated (i.e., the shared segments and non-shared segments). Now regarding the phonological selection, if it is sensitive only to the activation level of the target segments and does not consider the activation of non-target segments, then we anticipate a facilitation effect. However, if the phonological selection is sensitive to the activation level of both target and non-target segments, then we anticipate an interference effect. In other words, this will reverse the cognate effect; i.e., an interference effect will take the place of facilitation, due to the activated phonological representations of the non-target word being highly active and thus interfering with the selection of target segments. This interference effect would thus provide evidence that during word production, the phonological selection mechanism considers the activation level of the target and non-target segments. Figure 11 illustrates this hypothesis.

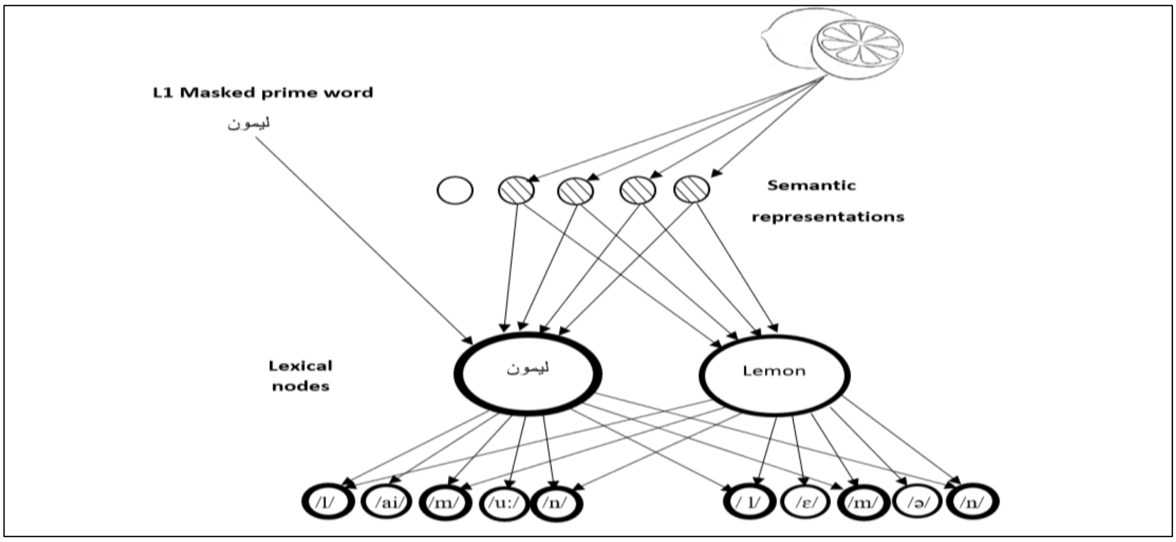


Figure. 11 A schematic representation of picture naming in L2 for cognate words words (/lɛmən/ and /laimu:n/) when primed by a masked L1 word. The arrows represent the flow of activation, while the thickness of the circles represents the level of activation of the representations.

The second objective of this experiment is to utilise the identity condition[14] to gain insights into the manner of lexical selection. The current debate is whether lexical selection is sensitive to the level of activation of both target and non-target lexical nodes, and thus entails competition (language non-specific view) or whether it is sensitive to only the target lexical nodes and there is no competition involved (language specific view). The findings of the identity effect in picture-word interference tasks indicates a language specific view of selection, as the facilitation effect was observed, instead of interference (Costa and Caramazza 1999). Therefore in experiment one, the participants were asked to name non-cognate pictures primed by the name of the picture in L1(masked). If only the target lexical node is considered for selection, more rapid naming latencies should be observed in the identity condition relative to the unrelated condition. However, if target and non-target lexical nodes are considered for selection, longer naming latencies should be found in the identity condition compared to the

[14] The identity condition refers to the condition during which the distractor word is the translation equivalent of the picture to be named in a picture interference task (Costa, Caramazza, et al.'s 2000).

unrelated condition. We also used non-cognate pictures to compare the reaction times in this condition to those in the cognates condition to determine whether the cognate pictures, despite the lack of orthographic overlap, would still be named faster than for non-cognates; details of this are provided next.

The third objective of this experiment was to test the effect of different script on lexical access and manner of selection. As discussed in sections 2.5 and 2.6, different script might influence lexical access and modulate the cross-language competition (e.g., Hoshino 2006; Miller and Kroll 2002). In experiment one, the prime words were written in Arabic script, which is linguistically different from English as reported in section 2.7. In addition, the task was monolingual, that is only English responses were required and the task instructions were given in English, in order to maintain the same mode. Thus, if different script modulates lexical access and selection, we expected the bilinguals in this study to ignore the prime word, and selectively access/select the target lexical nodes only. In other words, there would be no effect of prime type in the performance of the participants, as they would easily identify the prime being in the non-response language, and thus disregard it before processing it (see the language cue hypothesis discussed in section 2.6). If different script does not modulate lexical access, then we expected to find an effect of cross-language activation, i.e., either facilitation or interference. In addition, as discussed in section 2.5, that orthographic overlap is essential for obtaining the phonological facilitation effect in word production, which implies that orthography modulates phonological processing (Lupker 1982). Thus, we hypothesized that a similar orthography contributed to the cognate facilitation effect reported in same script studies (Costa, Caramazza, et al. 2000). If this theory holds true, then the performance of the different script participants would be the same when naming cognate and non-cognate pictures.

The fourth objective of this experiment was to identify whether the differences in the participants' proficiency level would have an impact on the nature of cross-language activation, the flow of activation, and the manner of lexical/phonological selection. Previous same script studies reported an enhanced speed and accuracy of lexical retrieval for highly proficient bilinguals, compared to less proficient bilinguals (e.g., Van Hell and Tanner 2012). Moreover, as discussed in section 2.4, the manner of lexical selection was argued to be modulated by the participants' proficiency level, namely highly proficient bilinguals applied language specific selection, whereas less proficient bilinguals applied language non-specific selection. However, the evidence available remains inconclusive, as most of these studies (e.g., identity condition: Costa and Caramazza 1999; cognate effect: Costa, Caramazza, et al. 2000) did not compare the performance of the two different proficiency groups in a single study. Therefore, here we aimed to fill this gap by testing two groups of bilinguals, the first of whom were being less proficient English-Arabic bilinguals, and the second highly proficient English-Arabic bilinguals. The next section provides a detailed account of the method used by the present study.

4.1 Method

4.1.1 Participants

As introduced in section 3.1, the research participants were seventy-six adult volunteers, all of whom were Saudi female Arabic-English bilinguals, for whom Arabic was their native language. All the participants had given consent to their data being used anonymously for research purposes prior to the study (see chapter three for further details about our ethics procedures). The data obtained from nine of the participants (of a total of seventy-six) was excluded from the analysis: six, due to technological malfunctions that caused no recorded responses to be retrievable, and two because they later explained to the researcher that they had been rushing and could not concentrate on the test, and the final participant was excluded for being trilingual (Arabic, English, and Indonesian), and was thus outside the scope of this bilingual trial.

In addition, all participants had taken the IELTS test. An independent-samples t-test was conducted to compare the IELTS scores for the two groups. This resulted in the identification of a statistically significant difference in proficiency level between the highly proficient group (M = 6.7, SD = .779) and the less proficient group (M = 3.8, SD = .84), $t(65) = -14.85$, $p < .001$. In the next section, the data acquired from the self-assessed rating for L2 proficiency level were analysed and reported.

4.1.1.1 The results of the language history questionnaire

As demonstrated in Table 2, the participants' average age at L2 acquisition was uniformly between ten and nine years old. All participants had primarily acquired English as a second language at an early age. In addition, all participants had received instruction in Arabic throughout their education, i.e., in their elementary, intermediate and high schools. The participants were asked to estimate their use of L1 and L2 on a daily basis. As shown in Table 2, the daily average estimation of the use of L1 and L2 by the less proficient group was 75.8% and 24.03%, respectively, while the average estimation for the highly proficient group for L1 and L2 usage daily was 44.44% and 55.56%, respectively.

The Highly Proficient Group (n= 36)	Minimum	Maximum	Mean	Std.Deviation
Age (years)	29	36	32.83	4.372
IELTS Score	6	8	6.7	0.779
Mean daily L1 usage (%) (5 pt scale)	25	100	55.56	18.039
Mean daily L2 usage (%)(5 pt scale)	25	100	44.44	18.039
Age of acquisition (years)	4	11	9.69	2.253

The Less Proficient Group (n= 31)	Minimum	Maximum	Mean	Std.Deviation
Age (years)	19	21	19.71	0.783
IELTS Score	3	4	3.8	0.832
Mean daily L1 usage (%) (5 pt scale)	50	100	75.81	16.437
Mean daily L2 usage (%)(5 pt scale)	< 25	50	24.03	16.453
Age of acquisition (years)	5	11	10.23	1.783

Table. 2 Descriptive statistics for the participants in experiment one.

All participants in the less proficient group had never experienced a long period of residence outside Saudi Arabia, i.e., in a location in which English was used as the primary mode of communication. All participants were residents of Saudi Arabia, aged between seventeen and twenty, who had been educated up to high school level.

Meanwhile, the highly proficient group (who were primarily teachers) were in possession of an MA. They were all resident in Saudi Arabia, but had spent some time living in the UK and/or the USA, where they had read for their Master's degree. Furthermore, many communicated with their colleagues in English during working hours, as these colleagues were non-Arabic speakers.

4.1.1.2 The results for the self-rating

An independent-samples t-test was conducted to compare the mean scores of the self-assessed rating for each skill between the two groups, leading to the identification of statistically significant differences between them in all skills areas: (i) Reading $t(65) = -8.75$, $p < .001$, (ii) writing $t(42.4) = -6.90$, $p < .000$, (iii) speaking $t(50.36) = -7.9$, $p < .000$, and listening $t(49.3) = -9.691$, $p < .000$. The results suggested that, based on their own perception of their proficiency level, the highly proficient bilinguals considered themselves to be native like, whereas the less proficient bilinguals considered themselves to be language learners. The results of the self-rating in the four skills are listed in Table 3.

Group Statistics	Proficiency	N	Mean	Std. Deviation	Std. Error Mean
Self-Rate Reading (5 pt scale)	High	36	4.72	0.513	0.086
	Less	31	2.9	0.7	0.126
Self-Rate writing (5 pt scale)	High	36	4.64	0.487	0.081
	Less	31	2.84	0.898	0.161
Self-Rate speaking (5 pt scale)	High	36	4.81	0.401	0.067
	Less	31	3.16	0.86	0.154
Self-Rate listening (5 pt scale)	High	36	4.94	0.232	0.039
	Less	31	3.03	0.983	0.176

Table. 3 Descriptive statistics for the self-assessed rating in participants' skills in L2 across the two groups

4.1.2 Materials

4.1.2.1 The selection of the target pictures

In total, sixty black and white line drawings of objects were selected for the stimuli list (further details in section 3.5.1). Six more pictures (with non-cognate names) were selected for the practice list and twenty for the filler list. Half the stimuli list had cognate names in Arabic and English, while the other remaining thirty were non-cognates (see appendix K for sample pictures). Each cognate picture name was carefully matched as closely as possible with a non-cognate picture name on the number of syllables and characters, frequency of words, imageability, familiarity and age of acquisition (see section 3.5.1) with which potential variables may impact picture naming latencies (Alario et al. 2004). The characteristics of the picture names used for the cognate and non-cognate conditions are summarised in Table 4.

An independent-samples t-test was conducted to ascertain whether any significant differences existed between the name of the cognate and the non-cognate pictures in English (means and results of *t*-test are reported in Table 4. The results revealed that the cognate and non-cognate pictures were matched in terms of the frequency of words, number of characters, imageability, familiarity, and the age of acquisition.

Variables	**Cognate Pictures Mean (SD)**	**Non-cognate Pictures Mean (SD)**	**t-test**	**p value (t-test)**
Name Agreement (%)	2.3 (1.6)	2.7 (2.1)	*t(58) = -1.006*	*p > .05*
Visual Complexity (KB)	18916.9 (10168.8)	15266 (6426.9)	*t(58) = .650*	*p > .05*
Syllable Length	2.3 (.75)	1.9 (.63)	*t(58) = 2.20*	*p < .05*
CharacterLength	6.0 (1.8)	5.9 (1.3)	*t(58) = .078*	*p > .05*
Frequency per million words	2.6 (1.5)	2.5 (1.1)	*t(58) = .276*	*p > .05*
Age of Acquisition (1-3 point scale)	2.1 (.899)	2.1 (.973)	*t(58) = .000*	*p > .05*
Imageability (100-700)	605 (31.02)	597 (22.6)	*t(58) = -.858*	*p > .05*
Familiarity (100-700)	518 (142)	499 (120)	*t(58) = 1.251*	*p > .05*

Table. 4 The characteristics of picture names in English that are used in experiment 1. Standard deviations are in parentheses.

The cognate words were longer in terms of their number of syllables than the non-cognate words, but we did not anticipate that this would be a problem as the critical comparisons would be made mainly between naming the cognate when it was preceded by related prime words vs.

unrelated prime words, and the same for the non-cognate pictures. Moreover, the speed/accuracy level of naming the cognates was not affected by their syllable length in the study by Hoshino et al. (2006), as the cognate facilitation effect was reported in their study, despite the cognate items having a greater number of syllables than the non-cognate items. In addition, a group of 33 Arabic-English balanced bilinguals were required to listen to audio recordings [15] and rate the phonological similarity of two types of sound pairs: (i) an English cognate name and its Arabic translation prime, and (ii) a non-cognate name and its Arabic translation prime. We employed a 5-point Likert scale, with "1" denoting very different and "5" very similar. The mean rating for each pair was then calculated and analysed. The results of the independent samples t-tests revealed significant differences in the rating scores between the cognate names paired with their L1 translation names, and the noncognate names paired with their L1 translation names [$t(58) = 23.11$, $p < 0.001$]. These cognate pairs were rated as phonologically similar more than the non-cognate pairs. The mean rating for cognate and non-cognate names is provided in appendix E.

4.1.2.2 The selection of prime words

All the pictures were preceded by prime words in the Arabic language, with two types of prime word being selected for each cognate picture, as follows: (i) a prime word forming the cognate name of the picture in L1 and (ii) a prime word unrelated to the name of the picture in L1. For example, a cognate picture of *a lemon* was primed with the cognate name in L1 (*ليمون* [English lemon]) in the related condition, and primed with an unrelated L1 word (*سياره* [English car]) in the unrelated condition. Two types of prime words were selected for each non-cognate picture: (i) a prime word in L1 consisting of the translation word for the name of the picture, and (ii) a

[15] The word pairs were recorded by a female balanced Arabic-English bilingual, who is an English language instructor at the English department, at Jeddah University. Following this, 33 Arabic-English postgraduate students at the English department were asked to rate these pairs.

prime word in L1 unrelated to the picture name. For example, a non-cognate picture of a bicycle was primed with the name of that picture in L1, i.e., the translation equivalent (دراجه [English bicycle]) in the related condition, and with an unrelated prime word in L1(عنب [English grapes]) in the unrelated condition.

Each related prime word was carefully matched with an unrelated prime word in terms of number of letters, number of syllables, and frequency (see section 3.5.2). Related and unrelated prime word lists were found to be similar in frequency, number of syllables, and number of characters (see Table 5 for these means) (all ps > .05).

Variables	**Related Primes Mean (SD)**	**Unrelated Primes Mean (SD)**	***t-test***	***p* value (t-test)**
Syllable Length	2.4 (.852)	2.3 (.800)	$t(118) = 1.215$	$p > 0.05$
Phoneme Length	5.3 (1.37)	5.0 (1.40)	$t(118) = 1.32$	$p > 0.05$
Written Form Frequency	2.6 (3.56)	2.7 (4.88)	$t(118) = .093$	$p > 0.05$

Table. 5 Characteristics of L1 prime words used in experiment 1. Standard deviations are in parenthesis.

4.1.2.3 The organization of material

Each participant was presented with sixty prime-target pairs: thirty of which were prime-target pairs for the cognate condition (i.e., fifteen related-cognate pairs and fifteen unrelated-cognate pairs); and thirty prime-target pairs for the noncognate condition (i.e., fifteen related-noncognate pairs and fifteen unrelated-noncognate pairs). To counterbalance the presentation of these pictures, two lists were created. The cognates pictures that were paired with related prime words in the first list were paired with unrelated prime words in the second list, and the same approach was applied for the noncognate pictures. Half the participants were presented with list 1 and the other half with list 2.

4.1.3 Design

The experiment was of a mixed design (i.e., a 2 x 2 x 2 design) with proficiency level (highly and less) representing the between-subjects factor, and the type of prime word (related and unrelated) and cognate status (cognate/non-cognate) of the pictures as within-subjects factors. Thus, all the participants in both the less proficient and the highly proficient group were asked to name pictures under the four conditions, as follows: (i) L1 related prime word preceding cognate pictures; (ii) L1 related prime word preceding non-cognate pictures; (iii) L1 unrelated prime word preceding cognate pictures; and (iv) L1 unrelated prime word preceding non-cognate pictures.

4.1.4 Procedures

The participants completed a six- question trial training session, which used different stimuli and prime words from those employed in the actual test. Then, they were directed to start the test when ready. As shown in Figure 12, each trial consisted of the following sequence of events: (i) the tests initially commenced with a fixation point (+), which appeared in the middle of the screen for 500 ms; (ii) the appearance of a visual mask of (#########) symbols replaced the fixation point, which remained for 500 ms; (iii) the appearance of a prime word (related or unrelated) appeared on screen for 50 ms; (iv) a visual mask of (#########) symbols, again appeared in the centre of the screen for 14 ms; and (v) a target picture, which appeared and remained on the screen until the participants responded. Further details regarding the experiment procedures are discussed in section 3.4. The next section presents an analysis of the reaction time and accuracy data.

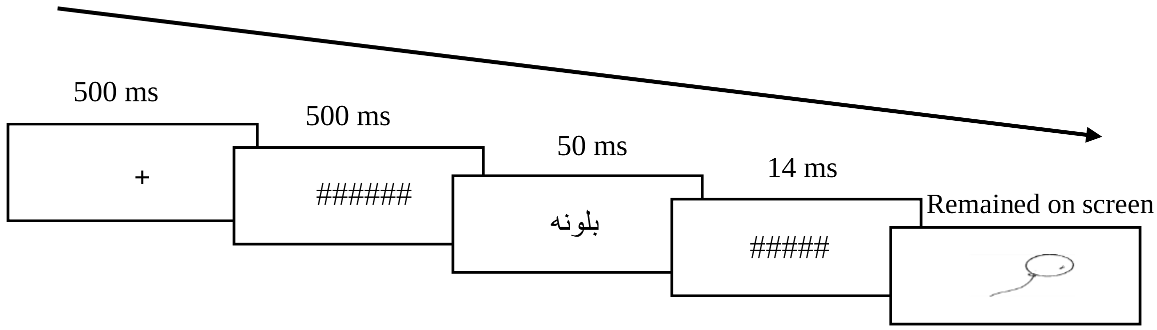


Figure. 12 The priming procedures applied in experiment 1. In this example, the picture of a balloon is primed with its cognate name in L1.

4.2 Analysis of Results

4.2.1 Reaction times analysis (RT)

The results of sixty-seven participants were included in the analyses. After trimming the data (as detailed in section 3.6.1), we then calculated reaction time and accuracy for correct responses again. Incorrect responses (5.80%) were excluded from the calculation. Errors and outlier trials were excluded from the following analysis:

(i) In the related cognate condition, errors formed 3.68% and the outliers formed 0.40%;

(ii) In the related non-cognate condition, errors formed 5.67% and outliers formed 0.90%;

(iii) In the unrelated cognate condition, errors formed 5.87% and outliers formed 0.50%;

(iv) In the unrelated non-cognate condition, errors at 7.96% and outliers at 1.79%.

The mean reaction times and percentage of accuracy rate are given in Table 6.

	Highly Proficient Group (n = 36)				Less Proficient Group (n = 31)			
	Prime Type				Prime Type			
Picture Type	Related Condition		Unrelated Condition		Related Condition		Unrelated Condition	
	RT	ACC	RT	ACC	RT	ACC	RT	ACC
Cognate Pictures	820 (64)	99.81 (0.11)	922 (83)	97.21 (4.3)	927 (81)	92.24 (6.2)	1051 (103)	90.54 (7.8)
Non-Cognate Pictures	1014 (133)	96.61 (4.8)	1181 (131)	95.52 (4.5)	1136 (105)	91.57 (7. 5)	1271 (87)	87.71 (7.7)

Table. 6 The mean response latencies (in ms) and accuracy of responses (in percentage) under the four conditions. Standard deviations are in parentheses.

A mixed ANOVA analysis was performed for the mean response latencies per subject, with: (i) proficiency level as a between-subjects factor, and (ii) the prime word type (related and unrelated prime words) and the cognate status of the pictures (cognate and non-cognate pictures) being within-subjects factors. Also, a mixed ANOVA was performed for the mean response latencies per items in each condition.

The results showed the main effect of the prime type was significant in the analysis by subjects $F1(1, 65) = 112.27$, MSE=1.16, $p < 0.001$, and in the analysis by items $F2$ (1, 58) = 241.38, $MSE = 1.02$, $p < 0.001$. The main effect of the cognate status of pictures was significant in the analysis by subjects $F1(1, 65) = 345.88$, $MSE = 3.23$, $p < 0.001$], and in the analysis by items $F2(1, 58) = 733.20$, $MSE = 2.91$, $p < 0.001$. Additionally, there was a significant main effect of proficiency level in the analysis by subjects $F1(1, 65) = 68.30$, $MSE = 0.83$, $p < 0.001$], and in the analysis by items $F2(1, 58) = 196.93$, $MSE = 0.75$, $p < 0.001$. The interaction between prime word type and proficiency level was insignificant in the analysis by subjects $F1(1, 65) = 0.05$, $MSE = 0.000$, $p = 0.830$ and in the analysis by items $F2(1, 58) = 0.11$, $MSE = 0.001$, $p = 0.744$. The interaction between the cognate status of the picture and proficiency level was insignificant in the analysis by subjects $F1(1, 65) = 0.26$, MSE =0.002, $p = 0.612$ and in the analysis by items *F2*(1, 58) = 0.51, *MSE* = 0.002, *p* =0.478. The interaction between prime type and picture type was insignificant in the analysis by subjects *F1*(1,65) = 2.49, *MSE* = 0.024, *p* = 0.119], and significant in the analysis by items *F2* (1,58) = 4.21, *MSE* = 0.018, *p* = 0.045. The interaction between prime type, picture type, and proficiency level was insignificant in the analysis by subjects *F1*(1,65) = 1.28, *MSE* = 0.012, *p* = 0.263, and in the analysis by items *F2*(1,58) = 1.76, *MSE* = 0.009, *p* = 0.190.

The results showed that the prime word type and cognate status of the pictures had a significant impact on the reaction time in general. The cognate effect and prime type effect were the same across the two proficiency groups, because the interaction between prime word type, proficiency level and the cognate status of pictures was insignificant (see Figure 13).

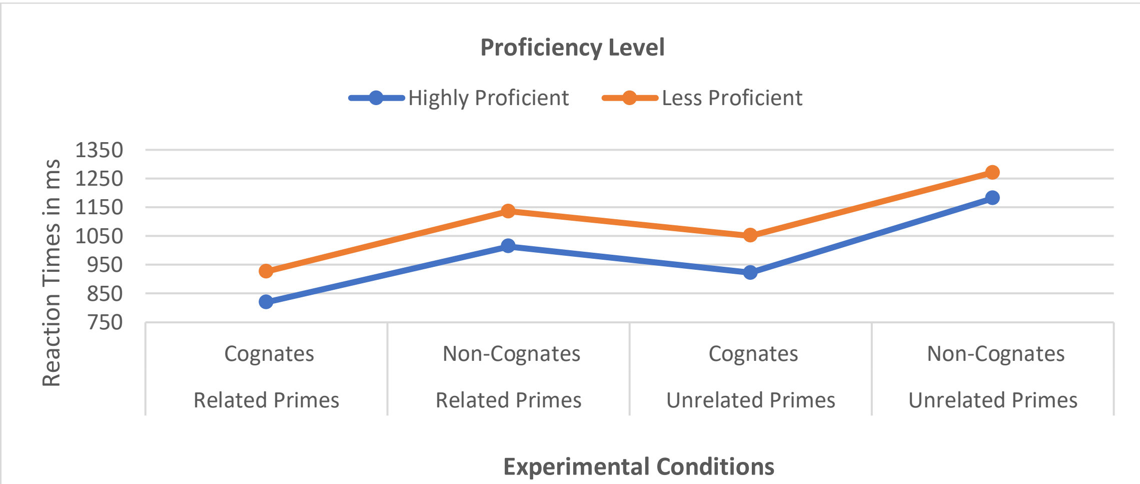


Figure. 13 The effects of prime word type and cognate status of pictures on reaction times across the high and less proficiency groups

As shown in Figure 13, the related prime words induced faster reaction times compared to the unrelated prime words for both proficiency groups. Moreover, the effect of prime type was the same for both the cognate and the non-cognate pictures, as the interaction between prime type and cognate status was insignificant. Moreover, cognate pictures were named significantly faster than the non-cognate pictures across all conditions for both the highly and less proficient groups. To examine the significance of the above noted differences, a paired sample t-test was conducted for each proficiency level. The results are given in Table 7.

Proficiency Level	Group 1	Group 2	Average RT (SD)		Difference between group 1 & group 2 (SE)	t-test (df)	p-value (Bonferroni correction)	Effect size (Cohen's d)[16]	Effect size tabulation
			Group 1	Group 2					
Highly proficient bilinguals	Related + Cognate	Related + Non-Cognate	820 (64)	1014 (133)	-194 (19)	- 9.74 (35)	< 0.001	- 1.62	large
	Related + Cognate	Unrelated + Cognate	820 (64)	922 (83)	-102 (15)	- 6.59 (35)	< 0.001	- 1.10	large
	Related + Cognate	Unrelated + Non-Cognate	820 (64)	1181 (131)	-361 (23)	- 15.3 (35)	< 0.001	- 2.55	large
	Related + Non-Cognate	Unrelated + Cognate	1014 (133)	922 (83)	92 (33)	3.59 (35)	= 0.003	0.60	large
	Related + Non-Cognate	Unrelated + Non-Cognate	1014 (133)	1181 (131)	-166 (29)	- 5.61 (35)	< 0.001	-0.94	large
	Unrelated + Cognate	Unrelated + Non-Cognate	922 (83)	1181 (131)	-259 (26)	- 9.78 (35)	< 0.001	- 1.63	large
Less proficient bilinguals	Related + Cognate	Related + Non-Cognate	927 (81)	1136 (105)	- 209 (22)	- 9.46 (30)	< 0.001	- 1.70	large
	Related + Cognate	Unrelated + Cognate	927 (81)	1051 (103)	- 124 (24)	- 4.97 (30)	< 0.001	-0.89	large
	Related + Cognate	Unrelated + Non-Cognate	927 (81)	1271 (87)	- 344 (24)	- 14.06 (30)	< 0.001	- 2.53	large
	Related + Non-Cognate	Unrelated + Cognate	1136 (105)	1051 (103)	85 (22)	3.78 (30)	= 0.002	0.68	large
	Related + Non-Cognate	Unrelated + Non-Cognate	1136 (105)	1271 (87)	- 135 (25)	- 5.25 (30)	< 0.001	-0.94	large
	Unrelated + Cognate	Unrelated + Non-Cognate	1051 (103)	1271 (87)	- 220 (26)	- 8.34 (30)	< 0.001	- 1.50	large

Table. 7 The results from the paired t- test to investigate the effect of prime word type and cognate status of pictures on the reaction times across the two proficiency groups.

[16] Cohen' d is one of the common ways to measure the effect size. It assesses the effect size of the differences between two means. According to Cohen (1988), $d = 0.2$ is considered a 'small' effect size, 0.5 represents a 'medium' effect size, and 0.8 a 'large' effect size (McLeod 2019).

Table 7 indicates that the fastest reaction times were naming cognate pictures preceded by related prime words, and the same result was observed for both the highly and less proficient participants. The differences in reaction times between cognate and non-cognate pictures in the related and unrelated conditions were statistically significant across both proficiency groups. This suggests the facilitation effect found in the cognate condition is not equal to that found in the non-cognate condition. Although unrelated prime words induced longer reaction times when they preceded cognate pictures, the reaction times under this condition were significantly faster than for the related and unrelated non-cognate conditions. This suggests that the cognates have privileged access compared to the non-cognates. As shown in Table 7, these differences proved to be statistically significant for both proficiency groups. For both proficiency groups, the longest reaction time was in the case of non-cognate pictures preceded by unrelated prime words. All the differences in reaction time for this condition and the other experimental conditions were statistically significant (Table 7).

As stated earlier, there is a significant main effect of proficiency level and as shown in Figure 13, the highly proficient group were faster at naming pictures than the less proficient group across all experimental conditions. Thus, we conducted a follow-up independent samples t-test to investigate the significance of the differences between the two proficiency groups across all four conditions (Table 8). As shown in Table 8, the average reaction time values for the highly proficient participants were significantly shorter compared to the average reaction time for the less proficient participants across all experiment conditions. The overall findings of reaction time analysis will be discussed in section 4.3.

Prime word type	Picture type	Average RT (SD) Highly proficient Group	Average RT (SD) Less proficient Group	Difference between highly & less proficient group (SE)	t- test (df)	p-value	Effect size Cohen's d	Effect size tabulation
Related	Cognate	820 (64)	927 (81)	- 106 (18)	- 5 .90 (57.24)	< 0.001	0.65	medium
Related	Non-cognate	1014 (133)	1136 (105)	- 122 (29)	- 4.18 (64.63)	< 0.001	0.45	medium
Unrelated	Cognate	922 (83)	1051 (103)	-128 (23)	- 5.54 (57.42)	< 0.001	0.61	medium
Unrelated	Non-cognate	1181 (131)	1271 (87)	- 89 (26)	- 3.23 (65)	0.002	0.35	small

Table. 8 The results for the independent samples t-test which investigates the differences in reaction time between the highly and less proficient groups across the experimental conditions.

4.2.2 Accuracy analysis (ACC)

We conducted a mixed ANOVA by participant with proficiency level (highly and less) as a between-subjects factor and prime word type (related and unrelated) and the cognate status of the picture (cognate and non-cognate) as within-subjects factors. In addition, we conducted a mixed-effect ANOVA by items with the cognate status of the pictures as a between-subjects factor, and proficiency level and prime word type as within-subjects factors.

The mixed ANOVA analysis showed a significant main effect of prime word type in the analysis by subjects $F1(1, 65) = 10.27$, $MSE = 357.77$, $p = 0.002$ and in the analysis by items $F2\ (1, 58) = 8.77$, $MSE = 349.85$, $p = 0.004$. The main effect of cognate status was significant in the analysis by subjects $F1\ (1, 65) = 8.55$, $MSE = 293.91$, $p = 0.005$ and in the analysis by items $F2\ (1, 58) = 5.84$, $MSE = 258.94$, $p = 0.019$. However, the interaction between the prime word type and proficiency level was insignificant in the analysis by subjects $F1\ (1, 65) = 0.41$, $MSE = 14.33$, $p = 0.522$ and in the analysis by items $F2(1, 58) = 0.44$, $MSE =14.55$, $p = 0.508$. This suggests the effect of prime type on accuracy was mostly the same for both proficiency groups and we discuss the implication of this finding in section 4.3 (see Figure 14). Moreover, there was an insignificant interaction between the cognate status of pictures and proficiency levels in the analysis by subjects $F1(1, 65) = 0.23$, $MSE = 8.01$, $p = 0.633$ and in the analysis by items $F2\ (1, 58) = 0.06$, $MSE = 2.26$, $p = 0.803$. This suggests the effect of the cognate status of the pictures on accuracy did not differ considerably between both proficiency groups (see Figure 14). Similarly, the interaction between the cognate status of pictures and prime type was insignificant in the analysis by subjects $F1(1, 65) = 0.04$, $MSE = 1.75$, $p = 0.838$, and in the analysis by items $F2(1, 58) = 0.09$, $MSE = 3.40$, $p = 0.771$. This suggests that, overall, the effect of prime type was the same for both cognate and non-cognate pictures.

Moreover, the three-way interaction between the cognate status of pictures, prime word type and proficiency was insignificant in the analysis by subjects $F1(1, 65) = 1.36$, $MSE = 55.88$, $p = 0.249$, and in the analysis items $F2(1, 58) = 1.94$, $MSE = 63.67$, $p = 0.169$. This suggests that the effect of prime type and cognate status of the pictures on accuracy did not differ significantly for either proficiency group.

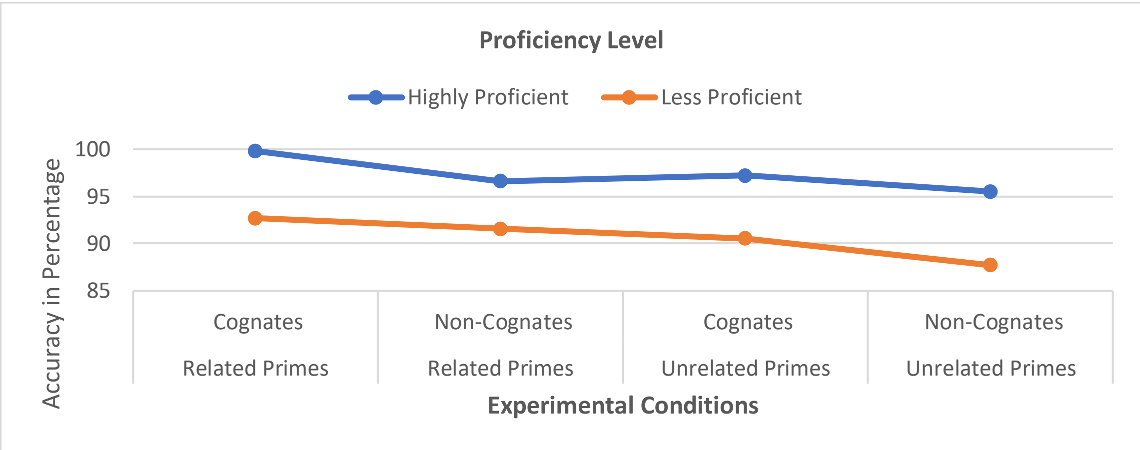


Figure. 14 The effect of prime type and cognate status of pictures on accuracy across the highly and less proficient groups.

As shown in Figure 14, the related prime words increased the accuracy rate for both proficiency groups when naming both the cognate and non-cognate pictures. However, as shown in Table

9, the analysis of group comparisons revealed that the differences were only significant when comparing the related-cognate condition to the other experimental conditions for the highly proficient group. The other differences noted in accuracy between related and unrelated prime conditions were not significant for the non-cognates for either the high or less proficient group.

Proficiency level	Group 1	Group 2	Average ACC (SD)		The difference between group 1 & group 2 (SE)	t-test (df)	p-value (Bonferroni correction)	Effect size (Cohen's d)	Effect size tabulation
			Group 1	Group 2					
Highly proficient bilinguals	Related+ Cognate	Related + Non-cognate	99.81 (0.11)	96.61 (4.8)	3.20 (0.80)	4.00 (35)	< 0.001	0.67	large
	Related Cognate	Unrelated + Cognate	99.81 (0.11)	97.21 (4.3)	2.61 (0.72)	3.63 (35)	0.003	0.60	large
	Related Cognate	Unrelated + Non-cognate	99.81 (0.11)	95.52 (4.5)	4.30 (0.81)	5.33 (35)	< 0.001	0.89	large
	Related + Non-cognate	Unrelated + Cognate	96.61 (4.8)	97.21 (4.3)	- 0.60 (1.07)	- 0.56 (35)	> 0.999	-0.09	small
	Related +Non-cognate	Unrelated + Non-cognate	96.61 (4.8)	95.52 (4.5)	1.10 (1.13)	0.97 (35)	> 0.999	0.16	small
	Unrelated Cognate	Unrelated Non-cognate	97.21 (4.3)	95.52 (4.5)	0.00 (0.00)	1.53 (35)	0.407	0.25	small
Less proficient bilinguals	Related Cognate	Related + Non-cognate	92.24 (6.2)	91.57 (7. 5)	0.68 (1.89)	0.36 (30)	> 0.999	0.06	small
	Related Cognate	Unrelated + cognate	92.24 (6.2)	90.54 (7.8)	1.71 (1.74)	0.98 (30)	> 0.999	0.18	small
	Related cognate	Unrelated + Non-cognate	92.24 (6.2)	87.71 (7.7)	4.54 (1.96)	2.31 (30)	0.084	0.42	small
	Related + Non-cognate	Unrelated + cognate	91.57 (7. 5)	90.54 (7.8)	1.03 (1.86)	0.55 (30)	> 0.999	0.10	small
	Related + Non - cognate	Unrelated + Non-cognate	91.57 (7. 5)	87.71 (7.7)	3.86 (2.27)	1.70 (30)	0.298	0.31	small
	Unrelated Cognate	Unrelated + Non-Cognate	90.54 (7.8)	87.71 (7.7)	2.83 (2.13)	1.33 (30)	0.581	0.24	small

Table. 9 The results of the paired t-test to investigate the effect of prime type and cognate status on accuracy rate across the two proficiency groups.

Moreover, the data in Table 9 indicates that the cognate status of the pictures had a significant impact on the performance of the highly proficient group only when the pictures were preceded by related prime words. The magnitude of the impact was insignificant for the highly proficient group in the unrelated condition. Conversely, for the less proficient group, the effect of cognate status of pictures on the accuracy rate was insignificant in both the related and unrelated conditions.

With regard to proficiency level, the ANOVA analysis showed the main effect of proficiency level was significant in the analysis by subjects $F1(1, 65) = 131.69$, $MSE = 3058.19$, $p < 0.001$, and in the analysis by items $F2(1, 58) = 59.74$, MSE = 1461.67, $p < 0.001$. The accuracy rate for the highly proficiency group was greater across all the conditions relative to the less proficient bilinguals (see Figure 14).

An independent sample t-test was conducted to calculate the size of the differences between the two proficiency groups under all conditions. As indicated in Table 10, the size of the difference between the two groups in terms of accuracy rate when naming those cognate pictures preceded by related primes was significantly large. Meanwhile, the difference in accuracy when naming non-cognate pictures in the related condition was small, and similar to the difference in accuracy rate for naming cognate pictures preceded by unrelated primes. Moreover, the difference between the two groups in terms of accuracy is of a medium size according to the calculated Cohen's test when naming non-cognate pictures preceded by unrelated prime words. All these findings will now be evaluated in the next discussion section (4.3) in light of how they address the research questions.

Prime Type	Picture Type	Average ACC (SD)		Difference between the High and less proficient groups (SE)	T- test (df)	P - value	Effect size (Cohen's d)	Effect size tabulation
		Highly proficient group	Less proficient group					
Related	Cognate	99.81 (0.11)	92.24 (6.2)	7.57 (1.13)	6.68 (31.65)	< 0.001	0.78	large
Related	Non-cognate	96.61 (4.8)	91.57 (7. 5)	5.05 (1.57)	3.22 (49.58)	0.002	0.36	small
Unrelated	Cognate	97.21 (4.3)	90.54 (7.8)	6.67 (1.52)	4.38 (65.00)	< 0.001	0.35	small
Unrelated	Non-cognate	95.52 (4.5)	87.71 (7.7)	7.81 (1.58)	4.95 (46.96)	< 0.001	0.56	medium

Table. 10 The results of the independent samples t-test to assess the differences in accuracy rate between the highly and less proficient groups across different conditions.

4.3 Discussion

This section first summarizes the findings and contributions of this study, and then compares them with those findings previously reported in the literature.

a) What is the manner of phonological selection/activation in different script bilinguals?

The results of this experiment revealed that cognate pictures in the related condition were named faster and more accurately than when preceded by unrelated prime words. The average difference in reaction time between the two conditions was 113 ms. The highly activated non-target segments, especially the non-shared ones, did not interfere with the selection process. This finding suggests that phonological selection considers only activated target segments, which contributes to answering the research question regarding the manner of phonological selection. The finding supports a language specific view of phonological selection. It also provides support to the view of Costa, Caramazza et al. (2000) that similar phonological segments in the non-target cognate name facilitate the selection of the target segments.

However, there is another possible explanation for the cause of the facilitation effect found in cognate naming that requires further investigation. It is possible that the phonological selection considers both the target and the non-target segments, but as the number of non-shared phonemes in the non-target node were relatively small compared to shared segments in this study, they were not strong enough to interfere. Even when it receives extra activation, as in this experiment, the non-shared phonemes do not cause interference. In order to investigate this hypothesis, a comparison should be made between two different types of cognates, i.e., cognates with high vs low phonological overlap. For example, the cognate name *ليمون* /laimu:n/ [English, lemon], has three overlapping phonemes (/l/, /m/, and /n/) and two different phonemes, whereas the cognate name *دولفين* /dolfi:n/ [English, dolphin] has five overlapping phonemes and one different phoneme. If the size of phonological overlap modulates selection, then we would expect to find naming cognates with a low phonological overlap (e.g., *lemon*) is not as fast as those with a high phonological overlap (e.g., dolphin). The manner of phonological selection should be noticeable when making within-cognates comparisons, rather than when comparing cognates to non-cognates.

In terms of this study's research questions regarding first the flow of activation, the findings of the cognate naming suggest a cascaded flow of activation, as it was evident that the non-target segments were activated at the phonological level, and that they facilitated the selection process. However, it is not possible to generalize this finding, as this cascaded pattern was found in the cognate condition only. The non-cognate pictures, which overlap in meaning only with the prime words, did not show any pattern of phonological activation, i.e., interference. If the non-target phonological segments were highly active, they would interfere with the phonological selection and cause interference. However, we found a facilitation effect that was not as equal as that identified in the cognate condition, which suggests a semantic overlap as the main cause of this effect. Thus, we argue that the findings suggest a cascaded flow of activation for cognate names only, and that phonological overlap is essential to allow a cascaded flow of activation. Secondly, the manner of phonological selection seems to be language specific in different script bilinguals, yet the view of non-specific selection cannot be ruled out completely. Moreover, this finding cannot be generalized as the cognate names have specific features that made them easily accessible in both languages, unlike the non-cognates. This was evident from the finding that even when preceded by unrelated primes, the cognate pictures were named faster and more accurately than the non-cognates in both the related and unrelated condition. Further research is therefore required to tap into the underlying processing of cognate naming; this lay outside the scope of the current study. However, the manner of phonological activation was examined further, but with non-cognate names, in the phoneme monitoring task in experiments four and five (chapters seven and eight, respectively).

The question of whether the dynamics of phonological activation are cascaded or discrete was investigated through various tasks. The finding of the cascaded view in the literature was the result of naming cognate pictures (Peterson and Savoy 1998; Janssen 1999) and phoneme monitoring tasks (Hermans et al. 1998; Colomé 2001). The results of this experiment confirm and expand upon the previous findings, as it demonstrates that the cognate facilitation effect, and the cascaded flow of activation are not limited to same script bilinguals, but are also the case for different script bilinguals. In terms of the manner of phonological selection, the results of this study confirm the typical findings of previous studies regarding naming cognates in picture word interference tasks and simple picture naming tasks (e.g., Costa, Caramazza, et al. 2000; Hoshino 2006).

b) What is the manner of lexical selection in different script bilinguals?

Regarding the manner of lexical selection, which is the second objective of this experiment, it was hypothesized that if the selection mechanism is sensitive to the activation level of both target and non-target lexical nodes in both languages (i.e., language non-specific view), an interference effect would be found in the identity condition; i.e., when participants named non-cognate pictures preceded by related primes. In contrast, if the selection mechanism is sensitive only to the level of activation of the target lexical nodes in the target language (i.e., language-specific view), a facilitation effect in the critical condition would be noted. This is because the non-target lexical node should be highly activated as it receives extra activation from the prime word (the L1 name of the picture), thus if selection is by competition, this highly active alternative should impede the selection process.

The results showed that naming non-cognate pictures in L2 produced a facilitation effect when preceded by masked related prime words in L1. Non-cognate pictures preceded by related primes were named faster than non-cognates preceded by unrelated primes, and the size of the difference between the two conditions was 150 ms. Furthermore, a higher accuracy rate was noted for the non-cognate pictures preceded by related primes, than for those preceded by unrelated primes.

The results indicate that both target and non-target lexical nodes were active at the lexical level, but only those in the target language were considered for selection. The observed facilitation effect could be attributed to the non-target lexical node sending extra activation to the target lexical node, as suggested by the language specific view. To illustrate, when participants named a picture of *a ring* preceded by the related prime word in the L1 (خاتم [English ring]), the non-target lexical node at the lexical level received activation from the picture and the prime word (خاتم [English ring]). The target lexical node received activation from the picture of a *ring* and the non-target lexical node at the lexical level. Being highly activated as it received activation from both the picture and non-target lexical item, the target lexical item was selected. The extra activation received from the non-target lexical item at the lexical level facilitated the production process. Thus, the non-target lexical node was not considered for selection.

Taken together, these findings are in accordance with the language-specific hypothesis, in that the two lexical items were activated at the lexical level but did not compete for selection. In other words, the language selection mechanism was not sensitive to the activation level of the non-target lexical nodes. In the following sections, these findings are explained in reference to previous studies on L2 word production.

In the psycholinguistic literature, the majority of support for language non-specific selection was derived from findings related to the semantic interference effect in the picture word interference tasks (Schriefers et al. 1990; Roelofs 1992; Starreveld and La Heij 1995; Hermans et al. 1998), while support for language-specific selection came from findings related to the identity effect in the picture word interference tasks (Costa and Caramazza 1999; Costa et al. 1999; Costa, Colomé, et al. 2000). As explained previously, the interference was interpreted to be a result of competition between the target and non-target lexical nodes. However, Costa et al. (1999) and Costa, Colomé, et al. (2000) argue that the semantic interference effect resulted from competition between lexical nodes in the target language (i.e., within language competition). For example, when a Spanish-English bilingual named a picture of a *dog* in English but was presented with a semantically related distractor word in the non-target language (e.g., the Spanish word *gato* [English cat], the semantic properties of '*gato*' would activate the corresponding lemma node in the target language '*cat*', which then might interfere with the selection of the target lexical node '*dog'*. Accordingly, it was argued that the semantic interference effect cannot adjudicate between the two views (Costa, Colomé, et al. 2000). In addition, the identity condition across the two languages was proposed as an alternative method when testing the predictions of the two views. Only a small number of previous studies tested this effect in same script bilinguals (e.g., Costa et al. 1999), and to the best of the author's knowledge, only one study examined different script bilinguals, namely that conducted by Hoshino (2006), who used a picture-word interference task and replicated the same facilitation effect. Thus, experiment one extends the findings of the identity effect in picture-word interference tasks by using a different method and a different group of bilinguals (Arabic-English speakers) whose two languages employ a distinct script.

Examining the two opposite views of lexical selection, we argue that the identity effect and the semantic interference effect can be traced back to different origins. Under the identity effect condition, the translation equivalent shares similar meanings (total overlap with the target picture name) but has a different orthography and phonological form; whereas, under the semantic interference condition, a semantically related word is used that partially shares the meaning with the target picture (e.g., orange and apple) but has completely different forms. At the conceptual level, it is possible for the presentation of a semantically related word to increase the number of activated concepts, unlike the presentation of the identical concept which is the case in the identity condition. Therefore, we posit that using prime or distractor words that have related concepts would lead to interference whereas using words with identical concepts would lead to facilitation. For example, we might present participants with three different pictures and ask them to name the picture in L2 if it includes an apple. In the task, picture A has a green and a red apple; B has an orange and an apple; and C has two unrelated objects. It is expected that picture A would elicit a faster response compared to B. Because to provide a correct response to picture B, the speaker has to untangle first the activated semantically related concepts ('*orange'* and '*apple'*) and then retrieve the correct answer, unlike picture A where the concepts are identical, and speakers must retrieve only the target name of the item.

Therefore, what is being argued here is that the semantic interference effect could be located at the conceptual level, and not at the lexical level. Thus, the long naming latencies observed under such conditions are not necessarily caused by competition at the lexical level but could be caused by competition at the conceptual level. This theory is addressed in the next chapters to help us address the research question regarding the manner of lexical selection by different script bilinguals. As a first step, we need to replicate the findings of the semantic interference effect found in the picture word interference tasks by using the masked priming task to exclude the possibility that this effect manifests only in such tasks due to the nature of the task's design, as explained in chapter two (section 2.3.2).

c) What is the effect of different script on lexical access and manner of selection?

The results of this experiment showed that L1 related prime words significantly affected the performance of the participants in the cognate and non-cognate conditions. This finding indicates that despite being presented in a different script, the Arabic prime words were processed and facilitated the production of the cognate and non-cognate names in the target language. Moreover, the different script did not act as a language cue to inhibit lexical access to the non-target lexical nodes. The non-target lexical nodes were activated at the lexical and phonological level, and this activation facilitated the production of the cognate names in the target language. These findings reflect those of the existing studies that tested different script bilinguals (Japanese-English: Hoshino 2006; Korean-English Moon and Jiang 2012; Arabic-French: Kheder and Kaan 2019), which reported different script did not inhibit lexical access to the non-target language. Moreover, Hoshino (2006) argued that a different script facilitated access to the target lexicon, as the performance of the different script bilinguals in their study was enhanced, compared to that of the same script bilinguals. The finding of experiment one in the present study is not informative in this matter, as we only tested different script bilinguals. Thus, further research is needed to validate this view across different scripts bilinguals.

Regarding the question of whether similar orthography is essential for obtaining the cognate facilitation effect in same script studies, the findings of the present study indicate that the cognate facilitation effect is robust, and that semantic and phonological overlap are sufficient to produce this effect. Therefore, finding a cognate facilitation effect by different script bilinguals indicates that the effect would manifest, despite script differences. The finding of this experiment concurs with that of Hoshino (2006), who found the presence of a cognate facilitation effect in Japanese-English bilinguals. Regarding the second research question concerning whether different script modulates lexical selection, the results of this study support the same pattern of findings reported in same script studies, in relation to language specific phonological selection (cognate naming: Costa, Caramazza, et al. 2000) and language specific lexical selection (identity condition: Costa and Caramazza 1999). However, the question of whether the manner of lexical/phonological selection entails competition was examined further using different experimental conditions by the current study (i.e., in experiments two, three, four, and five [chapters five, six, seven and eight, respectively]).

d) What is the effect of proficiency level on lexical access and manner of selection/activation?

The findings of cognates and non-cognates naming indicated that the different script bilinguals, specifically Arabic-English speakers, experience non-selective access that is similar to the same-script bilinguals, as discussed in the previous section (c). This section discusses whether or not the proficiency level of our participants modulated this cross-language activation pattern and lexical selection. In other words, as proficiency increases, do bilinguals selectively access the target lexicon and select the target lexical node? Also, do highly and less proficient bilinguals experience the same manner of cascaded flow?

The analysis of the data confirmed the presence of significant differences in the performance of the two groups under all the experimental conditions regarding reaction times and accuracy. In brief, proficiency level did affect retrieval time, as the highly proficient bilinguals named the pictures faster and more accurately than the less proficient group under all of the experimental conditions. This finding supports the assumptions of the Revised Hierarchical Model, as the speed and accuracy of lexical retrieval were affected by the bilinguals' proficiency level (Kroll and Stewart 1994). It also reflects same script studies that report delayed lexical activation and the retrieval of the weaker language (i.e., L2) in less proficient bilinguals compared to highly proficient bilinguals (e.g.,Van Hell and Tanner 2012). Therefore, the present study replicates the findings of previous same script studies, and provides evidence for the symmetrical effect of proficiency level on the performance of same and different script bilinguals during word production, in terms of speed and accuracy of retrieval/activation. Next, we discuss whether the two groups of bilinguals exhibit a different pattern of flow of activation and phonological selection.

An important finding concerning cognate naming was that both the less and highly proficient bilinguals experienced the same cognate facilitation effect, and that proficiency level did not impact the size of the cognate effect, as it was large for both the highly and less proficient bilinguals. This finding suggests that both proficiency groups experienced phonological activation and that the manner of phonological selection was language specific for both groups. Thus, to answer the research question regarding the effect of proficiency level on activation flow, and the manner of phonological selection, the findings suggest a cascaded flow and language specific phonological selection for both group of bilinguals. However, as discussed in section (a), this issue was investigated further in this study (experiments four and five [chapters seven and eight, respectively]), in order to confirm these findings and to address the research questions fully.

The findings of similar size of cognate effect for both proficiency groups is contrary to the claims made in previous studies (e.g., adult bilinguals: Costa, Caramazza, et al. 2000; bilingual children: Poarch and Van Hell 2012), which suggest that recognition of a cognate is affected by differences in proficiency level, and that the size of the effect is typically larger in one's weaker language. In the current experiment, a statistically considerable cognate effect size was identified for both the highly proficient and less proficient bilinguals, which suggested that proficiency level did not modulate the effect. The effect was noticeable for the less proficient bilinguals performing in their weaker language (i.e., L2) and also for the highly proficient bilinguals performing at near native ability in their L2. This experiment has the advantage of comparing two groups of bilinguals sampled form the same population (Arabic-English bilinguals) and using the same task, materials, procedures and response measurements. The arguments presented previously were based on studies that tested the effect of different proficiency level in L2 in naming cognates in L1 and L2, and found the size of the cognate effect in the L1 is smaller than in the L2 (Costa, Caramazza, et al. 2000). However, in this experiment we compared less and highly proficient bilinguals' performance in L2, which represents the weaker language for less proficient bilinguals, and relatively stronger language for highly proficient bilinguals. We strongly argue that when assessing the effect of proficiency level, comparisons be made between the performance of bilinguals in their second language, and not as a comparison between the mother tongue and the second language, because there are essential differences between L1 (the first language) and L2 (the second languages). For example, the acquisition of L1 differs from L2, as bilinguals are not only acquiring a means of communication, but also enhancing their knowledge of the world through their first language, unlike when acquiring their L2, which affords them a new way to talk about the world (Chenu and Jisa 2009). So, it is argued here is that judgments regarding the effect of proficiency level cannot be based on a comparison between first and second language, as the two languages differ essentially; no matter how proficient a speaker is in their L2, this cannot be compared to proficiency level in L1.

In his study, White (2015) addressed the same question, i.e., whether proficiency level modulates the cognate effect, by comparing the performance of two different proficiency groups (less and intermediate level) in a cognate naming task. She found a cognate effect occurred only in the group of less proficient bilinguals. However, there is a methodological flaw. The total number of participants was 22; half named the cognate pictures, and the other half named the non-cognate pictures, and thus formed a control group. So, five highly proficient participants in the cognate condition were compared to the performance of six highly proficient participants in the control condition and the same procedure was applied to the less proficient participants. We suggest that the sample size is too small to be representative. Also, Poarch and Van Hell (2012) compared the performance of less proficient children (e.g., 5th and 8th grade) with highly proficient adult bilinguals which is not a robust comparison as children have an underdeveloped L1 and L2 system. Experiment one addresses the gap in the literature by comparing the performance of adult bilinguals, and tests the effect of their different proficiency levels in the L2 in the same study, i.e., applying same experimental methods.

In terms of demonstrating an identity effect, the findings of the present study extend those of Costa and Caramazza (1999) and Costa et al. (1999), as facilitation was observed not only for the highly proficient bilinguals, but also for the less proficient bilinguals. This suggests that the manner of language selection is similar for both highly and less proficient bilinguals, and is not dynamic. This finding also adds to the existing literature regarding the manner of lexical selection by highly and less proficient bilinguals. The identify effect supports a language specific selection, however the previous investigations into this matter were conducted with highly proficient bilinguals. Thus, the present study bridges the gap in understanding by extending the findings to include less proficient bilinguals whose two languages have a different script. Finding a facilitation effect for both proficiency groups suggests that the manner of lexical activation is specific, however this view was further examined further in experiments two and three (chapters five and six, respectively).

This finding is contrary to reported findings in the literature. The question of whether proficiency level affects manner of lexical selection in bilinguals has been investigated primarily using language switching and mixing tasks (Meuter and Allport 1999; Jackson et al. 2001; Miller 2001; Costa and Santesteban 2004; Philipp et al. 2007; Schwieter and Sunderman 2008; Verhoef et al. 2010; Calabria et al. 2012), see sections 2.3.4 and 2.4.1 for a discussion of this matter. Such studies aimed to test the assumptions of the language non-specific selection view, that is the existence of an Inhibitory Control mechanism as proposed by (Green 1998). This control mechanism is thought to be responsible for controlling or suppressing the activation of the non-target lexical nodes at the lemma level. Green (1998) also assumed that it is harder to suppress the activation of the dominant language than the non-dominant language, because the dominant language is normally more active than the non-dominant language. The findings of these studies suggest inhibitory control tends to be found in the language production of unbalanced bilinguals (Jackson et al. 2001; Philipp et al. 2007; Schwieter and Sunderman 2008; Verhoef et al. 2010) but not in balanced bilinguals (Costa and Santesteban 2004; Calabria et al. 2012). This is because these speakers have reached a high level of proficiency and are able to apply a language-specific selection mechanism. Thus, previous findings suggest that the highly proficient bilinguals apply a language-specific selection mechanism, whereas less proficient bilinguals apply a non-specific language selection.

However, as discussed previously in the literature review chapter, (see section 2.3.4), there are limitation to the language-switching paradigm. According to Costa, La Heij, et al. (2006), this paradigm might clarify the control mechanism, but it is not informative regarding whether the non-response language is active during the speech production process. “This is because, arguably, in a language switching task participants may have their two languages active in a way that is not comparable to cases in which they are speaking in only one language” (Costa, La Heij, et al. 2006, p. 141). The task artificially stimulates cross language activation by requiring participants to name target words or pictures in two languages. The procedures and the stimuli force the participants to adopt a bilingual mode, which according to Grosjean (2001) could result in the nonselective activation reported in these studies. Therefore, the findings of these studies might not be adequate to describe the normal/natural production process. In contrast, we argue that the methodology applied in this study provides a good simulation of the natural process of single word production. The participants were not forced to use a bilingual mode as they were only asked to name the pictures in their L2. Additionally, they were not informed about the prime words. The instructions given were to name the pictures as fast and as accurately as possible. In summary, the findings of this experiment suggest that highly and less proficient bilinguals activate both lexicons (L1 and L2) when naming an object in L2, and that proficiency level does not modulate the cross-language activation.

4.4 Conclusion

In this experiment, the cognate effect and the identity effect were used to explore the issue of lexical access and selection using the priming method instead of the traditional method of picture word interference tasks, in order to address our research questions concerning the manner of lexical access, the flow of activation, and lexical/phonological selection in different script bilinguals. It also addressed the matter of whether a different script modulates the manner of activation/selection. In addition, we investigated the effect of proficiency level, and whether it modulates lexical access and the manner of lexical/phonological selection. We concluded that different script bilinguals (i.e., Arabic-English speakers) experienced non-selective lexical access when naming pictures in their L2. Moreover, the flow of activation was cascaded when naming cognate pictures (i.e., at the phonological level). Regarding the language selection mechanism, we found that the lexical and phonological selection mechanism was language specific for both proficiency groups. These findings extended the existing knowledge regarding lexical access and manner of selection to include different script bilinguals. Furthermore, the experiment provided evidence that the identity and cognate effect is robust, even when tested in a different paradigm, i.e., the masked priming paradigm. However, we concluded that further research is essential to establish whether the contradiction of findings in the literature was due to differences in the locus of the semantic effect and the identity effect. Finally, we concluded that proficiency level did not modulate cross-language activation or the manner of lexical/phonological selection. We investigated less proficient bilinguals, a group who received little attention previously, and compared their performance to a highly proficient group in different critical conditions, i.e., identity and cognate naming. We provided a preliminary analysis of the manner of selection and activation by this group, which the findings of experiment one demonstrated was like that of highly proficient bilinguals. Further comparisons of the performance of these two proficiency groups were conducted via different experiments included in this study (chapters five, six, seven, and eight), in order to gain a comprehensive overview of the production process.

Chapter 5 Experiment Two: Semantically related Primes in a Picture Naming task

In experiment one, it was evident that priming a non-cognate picture with its translation equivalent in L1 yielded a facilitation effect consistent with the language-specific view in that non-target lexical nodes were not considered for selection. This finding contrasted with several studies in the literature (e.g., Schriefers et al. 1990; Roelofs 1992; Starreveld and La Heij 1995), which argued in favour of the non-specific selection view due to the semantic interference effect. These studies mainly used the picture-word interference paradigm. Counter arguments attributed the findings associated with the semantic interference effect to the nature of the task. More specifically, they (e.g., Kroll et al. 2005) argued that the presence of a distractor word was easily detectable by participants, who might formulate a (covert) verbal response to the distractor word, which would become available for production prior to the picture-naming response, resulting in the observed interference effect (Finkbeiner and Caramazza 2006, p. 790). Therefore, in experiment two, detailed here, we tested whether priming a picture with a semantically related masked prime would yield an interference effect. The advantage of this paradigm was that the semantically related prime was masked and presented briefly, so as not to be visually detected nor consciously processed, thereby ruling out any task related factors, and enabling us to observe how a semantically related prime affects the cross-language activation process. Thus, the primary objective of experiment two was to test whether we can replicate the reported findings concerning the semantic interference effect by using a different paradigm. More specifically, finding semantic interference would imply: (i) there is parallel activation across the two languages, and (ii) the manner of language selection at the lexical level may entail competition.

In experiment one, non-cognate pictures produced a facilitation effect when preceded by a translation equivalent at stimulus onset asynchrony (SOA) of 64 ms. Here in experiment two, we wanted to establish if the non-cognate pictures would produce interference when preceded by semantically related primes. Semantic interference effects were generally found, in picture word interference tasks when the distractor word was presented before the target pictures (Starreveld and la Heij 1996). Also, in the two masked priming studies with monolinguals, the effect was found at 100 ms (Alario et al. 2000; Bajo et al. 2003). Thus, in this experiment we varied the time of the presentation of the masked prime to 50, 75 and 100 ms to determine whether any semantic interference effect, if found, would vary in different script bilinguals with distinct proficiency levels.

Regarding the question of whether script differences modulate lexical access and cross-language activation, the findings of experiment one suggest that Arabic-English bilinguals experience non-selective access, and that script differences do not affect the manner of selection/activation. This issue was examined further in this current experiment and the next three experiments (three, four, and five [chapters six, seven, and eight, respectively]). Before making any generalization, it was necessary to determine whether non-selective access and cross-language activation persisted, despite the manipulations carried out in these experiments. Thus, we hypothesized that if different script acts as a language cue, then related prime words should not affect the speed and accuracy of retrieval. This is because they were written in the non-response language (i.e., Arabic), and the participants were required to respond in English. As there are no similarities between the Arabic and the English language in the written form (see section 2.7), the participants would be able to ignore or inhibit processing the prime words easily. Moreover, finding a semantic interference effect that is typically reported in same script studies would suggest that script differences do not modulate the manner of lexical selection.

Also, in this experiment we investigated whether proficiency level would modulate cross-language activation. In experiment one, when we primed cognate and non-cognate pictures with related prime words, we found no effect from proficiency level on either cross-language activation, or the selection process. To investigate whether this finding would hold when priming target pictures with semantically related words, we tested two groups of adult Arabic-English bilinguals (highly and less proficient). If proficiency level modulates cross-language activation, then we would expect the size of the semantic interference effect to be dependent on the participants' proficiency level. The following section presents a detailed account of the methods used in this experiment, followed by an analysis of the results and a discussion section.

5.1 Method

5.1.1 Participants

The research participants were sixty-three new (not previously recruited) adult volunteers, all of whom were Arabic-English bilinguals with Arabic as their native language. All participants had given consent to their data being anonymously used for research purposes before joining the study (see section 3.2 for further details). The participants were divided into two groups according to their proficiency level (highly and less) as detailed in chapter 3, section 3.1. The data from three participants (from a total of sixty-three) was excluded from the analysis due to technological malfunctions, whereby no recorded responses were collected. That left sixty participants remaining.

An independent-samples t-test was conducted to compare the IELTS scores of the two groups. This resulted in the identification of a statistically significant difference between the highly proficient group (M =7.4, SD=1.08) and the less proficient group (M = 3.73, SD =.783), t (52.8) = -15.320, $p < .001$, in terms of proficiency.

5.1.1.1 The results of the language history questionnaire

As demonstrated in Table 11, the participants' average age of acquisition of L2 was uniformly between nine and ten years of age. All participants had primarily acquired English as a second language through formal classroom teaching. In addition, all had received instruction in Arabic, i.e., at their elementary, intermediate and high schools. The participants were asked to estimate their daily use of L1 and L2. As shown in Table 11, the daily average estimation of the use of L1 and L2 by the less proficient group was 76% and 23%, respectively, while the average estimations for the highly proficient group of L1 and L2 usage per day was 48% and 51%, respectively. Refer to appendix C to see a copy of the language history questionnaire.

The Highly Proficient Group (n= 30)	**Minimum**	**Maximum**	**Mean**	**Std.Deviation**
Age (years)	27	36	31.23	2.3
IELTS	6	9	7.43	1.08
Mean daily L1 usage (%) (5 pt scale)	25	75	48.67	16.23
Mean daily L2 usage (%) (5 pt scale)	25	75	51.33	16.13
Age of acquisition (years)	4	11	9.43	2.38
The Less Proficient Group (n=30)	**Minimum**	**Maximum**	**Mean**	**Std.Deviation**
Age (years)	19	22	20	0.794
IELTS	3	5	3.7	0.783
Mean daily L1 usage (%) (5 pt scale)	50	100	76.6	16.47
Mean daily L2 usage (%) (5 pt scale)	< 25	50	23.17	16.47
Age of acquisition (years)	5	11	10.2	1.8

Table. 11 Descriptive statistics for participants in experiment two.

5.1.1.2 The results for the self-rating:

An independent-samples t-test was conducted to compare the mean scores for the self-assessed rating in each skill between the two groups, leading to the identification of statistically significant differences between the two groups across all skills: (i) Reading $t(32.6) = 15.28$, $p < .001$, (ii) writing $t(58) = 16.15$, $p < .001$, (iii) speaking $t(46) = 15.41$, $p < .001$, and listening $t(29) = 14.32$, $p < .001$. The results suggested that, based on their own self-perceived proficiency level, the highly proficient bilinguals considered themselves as native like, whereas the less proficient bilinguals considered themselves to have achieved a basic level of proficiency level. Table 12 presents the descriptive statistics for the self-assessed rating.

Group Statistics	Proficiency Level	N	Mean	Std. Deviation	Std. Error Mean
Self-Rate Reading (5 pt scale)	High	30	4.97	0.183	0.033
	Less	30	2.87	0.73	0.133
Self-Rate writing (5 pt scale)	High	30	4.67	0.479	0.088
	Less	30	2.47	0.571	0.104
Self-Rate speaking (5 pt scale)	High	30	4.87	0.346	0.063
	Less	30	2.9	0.607	0.111
Self-Rate listening (5 pt scale)	High	30	5	0	0
	Less	30	2.9	0.803	0.147

Table. 12 Descriptive statistics for the self-assessed rating in participants' skills in L2 across the two groups.

5.1.2 Materials

5.1.2.1 The selection of the pictures

Sixty-five words were adopted from the Prototypically norms for 26 Semantic Category (Uyeda and Mandler 1980). Thirty for the stimuli list, twenty for the filler list, and six for the practice list (see appendix F for a complete list of pictures). Then the corresponding black and white line-drawings were adopted from the International Picture Naming Project database as detailed in section 3.5.1(see appendix K for sample pictures). All picture names were non-cognates and some were included in the non-cognate condition in experiment one. However, this did not affect the result of this experiment, as a different group of participants took part in this study. All selected pictures have high imaginability, frequency, and familiarity rating.

5.1.2.2 The selection of the prime words

For each of the target pictures, two types of prime words were selected: (i) a semantically related prime word (e.g., *fork -spoon*), and (ii) an unrelated prime word from a different category (e.g., *fork-monkey*). Similar to pictures names, eighty-six prime words were selected from the Prototypically norms for 26 Semantic Category (Uyeda and Mandler 1980): (i) sixty related and unrelated primes for picture stimuli; (i) twenty unrelated primes for fillers; and (iii) six unrelated primes for practice list.

A group of 50 Arabic-English bilinguals were requested to rate the two lists to establish their similarity on a 5-point Likert scale, with “1” denoting very different and “5” very similar. The mean rating for each prime-target word pair was calculated and analysed. The results of the paired samples t-tests revealed a significant differences in the rating scores, [$t(29) = 43.2$, $p < 0.001$]. The semantically related list was rated as more similar than the unrelated list. The mean rating for each list (semantically related pairs and unrelated pairs) is summarized below in Table 13.

Conditions	Mean	SD
Semantically Related Pairs (n 30)	4.3	.297
Unrelated Pairs (n 30)	1.3	.301

Table. 13 Similarity rating for semantically related prime-target picture pairs and unrelated prime-target picture pairs.

We ensured that each related prime word was matched as closely as possible with an unrelated prime word on number of letters, number of syllables, and frequency (see section 3.5.2). Related and unrelated prime word lists were found to be similar in frequency, number of syllables, and number of characters (see Table 14 for these means) (all ps > .05).

Variables	Related Primes M (SD)	Unrelated Primes M (SD)	t-test	p value (t-test)
Syllable Length	2.23 (0.89)	2.1 (.77)	*t(58) = .95*	*p > 0.05*
Phoneme Length	5.4 (1.6)	5.0 (1.3)	*t(58) = 1.106*	*p > 0.05*
Written Form Frequency	9.8 (32.0)	10.5 (13.4)	*t(58) = .108*	*p > 0.05*

Table. 14 Characteristics of L1 prime words used in experiment two.

1.1.1.1 The organization of material

In experiment two, each participant was presented with 30 prime-target word pairs: fifteen were prime-target picture pairs for the semantically related condition (set A), and fifteen prime-target picture pairs for the unrelated condition (set B). The prime-target picture pairs were counterbalanced, resulting in the creation of two lists. We matched picture names (across conditions) on variables that are often considered critical to the type of processing level involved in the task and experimental conditions (see section 3.5.1). The picture names were controlled at ps > .05 (see Table 15 for these means).

Group Statistics	set A (N=15)	set B (N= 15)	*t-test*	***p* value (t-test)**
Name Agreement (%)	0.96 (.076)	0.92 (0.158)	*t(28) =.869*	*p > .05*
Visual Complexity (KB)	17646.4 (8596.6)	17923.07 (8269.6)	*t(28) = -.090*	*p > .05*
Syllable Length	1.4 (.632)	1.6 (.828)	*t(28) = -.743*	*p > .05*
CharacterLength	4.93 (1.831)	5.07 (1.831)	*t(28) = -.199*	*p > .05*
Frequency per million words	3.47 (1.163)	3.37 (1.437)	*t(28) =.215*	*p > .05*
Age of Acquisition (1-3 point scale)	1.33 (.617)	1.47 (.743)	*t(28) = -.535*	*p > .05*
Imageability (100-700)	546.8 (43.1)	551.47 (60.1)	*t(25.3) = -.244*	*p > .05*
Familiarity (100-700)	615.27 (17.74)	611.93 (35)	*t(28) =.329*	*p > .05*

Table. 15 Characteristics of pictures and pictures' names used in experiment two.

Pictures that were paired with related prime words in the first list were paired with unrelated prime words in the second list. Since there were three different presentation times (50, 75, and 100 ms), we ensured that each list was presented for 50, 75, and 100 ms. Therefore, a total of six task files were created, and divided between the participants.

5.1.3 Design

The experiment was of a mixed design (i.e., a 2 x 2 x 3 design) with proficiency level and the presentation time of the prime words (50, 75, and 100 ms) representing the between-subjects factor and the type of prime (i.e., related and unrelated) as within-subjects factor. Thus, all participants in both the less proficient and the highly proficient group were requested to name pictures preceded by masked prime words. Overall, there were six experimental conditions: (i) semantically related primes presented for 50 ms; (ii) unrelated primes presented for 50 ms; (iii) semantically related primes presented for 75 ms; (iv) unrelated primes presented for 75 ms; (v) semantically related primes presented for 100 ms; and (vi) unrelated primes presented for 100 ms.

5.1.4 Procedures

The procedure was identical to that used in Experiment one, with the exception that the presentation time for the prime words varied i.e., primes were presented for 50 ms, 75 ms and 100 ms. The interstimulus interval (ISI) was fixed to 14 ms. Hence, in this experiment, three Stimulus-Onset-Asynchronies were used: 64, 89 and 114 ms, respectively. See methodology (chapter 3) for further details.

5.2 Analysis of Results

5.2.1 Reaction times analysis (RT)

The data were collected from 60 participants. After trimming the data (detailed in section 3.6.1), we calculated the reaction times and accuracy for correct responses again and removed errors and outliers. Table 16 lists the percentages of errors and outliers that were removed.

Prime Type	Related Primes	Unrelated primes	Related primes	Unrelated primes	Related primes	Unrelated primes
Presentation times	50 ms	50 ms	75 ms	75 ms	100 ms	100 ms
Errors	7.0 %	6.6 %	4.33 %	5.33 %	21.0 %	7.0 %
Outliers	6.2 %	1.0 %	1.33 %	1.33 %	1.67 %	0.67 %

Table. 16 The percentage of errors and outliers removed from the analysis in the six conditions.

We calculated reaction times and accuracy for correct responses again. The mean reaction times and percentage for accuracy rate are shown in Table 17.

Picture Type	High Proficiency (n = 30)				Low Proficiency (n = 30)			
	Prime Type				Prime Type			
	Related Condition		Unrelated Condition		Related Condition		Unrelated Condition	
	RT	ACC	RT	ACC	RT	ACC	RT	ACC
50 ms	1240 (86)	91.90 (6.8)	1214 (73)	97.33 (6.4)	1363 (94)	93.95 (3.7)	1330 (87)	91.90 (8.1)
75 ms	1286 (82)	96.62 (4.7)	1218 (104)	94 (6.6)	1426 (66)	94.67 (6.1)	1384 (102)	92.67 (5.8)
100 ms	1391 (108)	81.81 (9.4)	1214 (84)	95.33 (5.4)	1555 (86)	75.52 (9.4)	1306 (78)	90.56 (13.01)

Table. 17 Mean response latencies (in ms) and the accuracy rate of responses (in %) across all experimental conditions. Standard deviations are in parentheses.

A mixed ANOVA analysis was performed on the mean response latencies per subject with: (i) proficiency level and prime presentation times (50, 75 and 100 ms) as between-subjects factors, and (ii) prime type (related and unrelated prime words) as a within-subjects factor. In addition, we conducted a repeated measure ANOVA on the mean response latencies per items, with prime word type, presentation times, and proficiency level as within-subjects factors.

The results showed a significant main effect of prime type in the analysis by subjects $F1(1, 54) = 56.46$, $MSE = 0.30$, $p < 0.001$, and in the analysis by items $F2\ (1, 29) = 53.06$, $MSE = 0.84$, $p < 0.001$. Moreover, a significant main effect was found for the presentation time of prime words in the analysis by subjects $F1\ (2,54) = 5.95$, $MSE = 0.06$, $p = 0.005$, and in the analysis by items $F2\ (2,29) = 8.90$, $MSE = 0.18$, $p < 0.001$. There was a significant main effect of proficiency in the analysis by subjects $F1\ (1, 54) = 50.52$, $MSE = 0.54$, $p < 0.001$, and in the analysis by items $F2\ (1, 29)=73.98$, $MSE = 1.59$, $p < 0.001$. Also, a significant interaction was found between the type of the prime word and prime presentation times in the analysis by subjects *F1*(2, 54) = 18.86, *MSE* = 0.10, $p < 0.001$, and in the analysis by items *F2*(2, 58) = 12.09, *MSE* = 0.30, $p < 0.001$.

This interaction indicates that the effect of prime type on reaction times was dependent on the presentation time. Longer presentation times caused longer reaction times in the related condition relative to the unrelated condition. Figure 15 shows that as the presentation time increased from 50 to 75 ms, reaction times increased in the related condition, but insignificantly. However, further increasing to 100 ms led to a significant increase in reaction times. This difference will be explained in the discussion section 5.3.

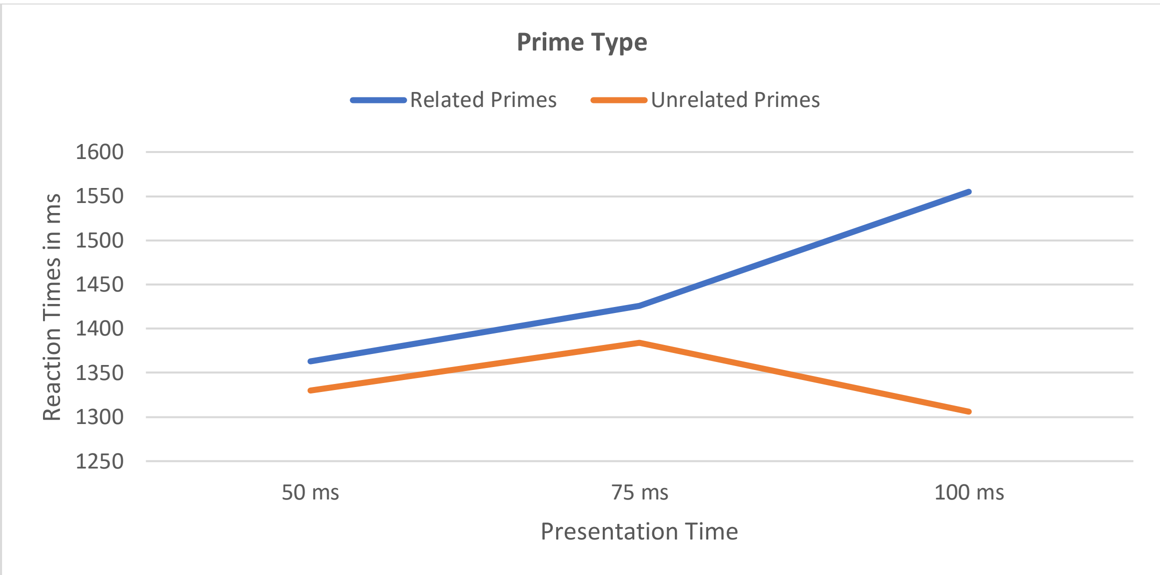


Figure. 15 The effects of presentation times on reaction times across the related and unrelated conditions.

There was an insignificant interaction between proficiency level and the type of prime words in the analysis by subjects *F1*(1, 54) = 0.45, *MSE* = 0.002, $p = 0.50$ and in the analysis by items *F2*(1, 29) = 0.68, *MSE* = 0.01, $p = 0.417$. In addition, the interaction between proficiency level and presentation times was insignificant in the analysis by subjects *F1(*1,54) = 0.29, *MSE* =

0.03, *p* = 0.119, and in the analysis by items *F2* (2, 58) = 0.23, *MSE* = 0.01, *p* = 0.797. Moreover, the interaction between the type of prime words, presentation time and proficiency level was insignificant in both the analysis by subjects *F1* (2,54) = 1.21, *MSE* = 0.01, *p* = 0.305, and in the analysis by items *F2* (2, 58) = 0.36, *MSE* = 0.01, *p* = 0.701.

Having said that, the type of prime words, presentation time, and proficiency level had a significant impact on reaction times in general, and the effects were the same across the two different proficiency groups, because there was an insignificant three-way interaction as reported herein.

A post-hoc test was conducted to check the significance of the noted differences in reaction times between the three different types of presentation times under the related and unrelated prime word conditions (Table 18).

Type of prime word	Group 1	Group 2	Average RT (SD)		Difference between Group 1 & Group 2 (SE)	T-test (df)	p-value (Bonferroni correction)	Effect size (Cohen's d)	Effect size tabulation
			Group 1	Group 2					
Related	50 ms	75 ms	1302 (108)	1356 (102)	-54 (33)	-1.62 (38)	0.226	0.263	small
	75 ms	100 ms	1356 (102)	1473 (127)	-117 (36)	-3.19 (38)	0.006	0.517	medium
	50 ms	100 ms	1302 (108)	1473 (127)	-171 (37)	-4.57 (38)	0.000	0.741	medium
Unrelated	50 ms	75 ms	1272 (98)	1301 (131)	-29 (36)	-0.78 (38)	0.878	0.127	small
	75 ms	100 ms	1301 (131)	1260 (92)	41 (36)	1.14 (38)	0.524	0.185	small
	50 ms	100 ms	1272 (98)	1260 (92)	12 (30)	0.40 (38)	> 0.999	0.065	small

Table. 18 Independent sample t-test results of the effect of duration time on RT.

The results indicated that the differences in reaction times between the three different presentation times in the unrelated condition were not significant. In addition, the differences in reaction times between the two conditions (50 ms and 75 ms) was insignificant in the related condition. However, when comparing reaction times in the 100 ms condition with the other two conditions (i.e., 50 ms and 70 ms), the difference reached significance level. The same pattern was noticed for both proficiency groups, as the interaction between prime types and proficiency level was insignificant.

We conducted a paired t-test to examine the significance of the differences in reaction times between related and unrelated distractors at each presentation time (see Table 19).

Presentation times	Average RT (SD)		Difference between related & unrelated (SE)	T-test (df)	P - value	Effect size (Cohen's d)	Effect size tabulation
	Related prime words	Unrelated prime words					
50 ms	1302 (108)	1272 (98)	30 (26)	1.13 (19)	0.272	0.260	small
75 ms	1356 (102)	1301 (131)	54 (22)	2.47 (19)	0.023	0.567	medium
100 ms	1473 (127)	1260 (92)	212 (20)	10.89 (19)	0.000	2.497	large

Table. 19 The results of the paired t-test to investigate the effect of the prime word type and different presentation times on reaction times.

According to Table 19, the effect of prime type was insignificant on reaction times when the presentation times were 50 ms. However, the related prime words presented for 75 ms and 100 ms induced statistically significant longer reaction times, relative to the unrelated prime words. This pattern was evident in both proficiency groups, because the three way interaction between prime type, presentation and proficiency level was insignificant.

Further, as reported earlier in this section, proficiency level had a significant effect on reaction times. We conducted an independent sample t-test to compare the differences in average reaction times between the highly and less proficient participants under related and unrelated prime word conditions (Table 20).

Prime word type	Average RT (SD)		Difference between the two groups (SE)	T-test (df)	P - value	Effect size (Cohen's d)	Effect size tabulation
	Highly proficient	Less proficient					
Related	1306 (110)	1448 (114)	-0142 (29)	-4.91 (58)	< 0.001	0.644	large
Unrelated	1215 (85)	1340 (93)	-124 (23)	-5.41 (58)	< 0.001	0.710	large

Table. 20 The results of the independent sample t-test to investigate the effect of proficiency level on reaction times for different prime word types.

As shown in Table 20, the reaction times for the highly proficient participants were shorter on average than for the less proficient participants where there were both related and unrelated distractors. This pattern proved to be the same for all durations, because the interaction between presentation times and proficiency level was insignificant.

An independent t-test was conducted to see the effect of proficiency level across all experimental conditions (Table 21). According to Table 21, the size of difference between the highly proficient and less proficient bilinguals was significantly larger in the related condition when presentation times were 75 ms and 100 ms. More specifically, the impact of the related primes in these two conditions on the less proficient participants was greater than that on the highly proficient bilinguals.

Presentation times	Prime word type	Average RT (SD)		Difference between high & less (SE)	T-test (df)	p-value	Effect size (Cohen's d)	Effect size tabulation
		Highly proficient	Less proficient					
50 ms	Related	1240 (86)	1363 (94)	-122 (40)	-3.02 (18)	0.007	0.713	middle
	Unrelated	1214 (73)	1330 (87)	-116 (36)	-3.23 (18)	0.005	0.761	middle
75 ms	Related	1286 (82)	1426 (66)	-140 (33)	-4.17 (18)	0.001	0.984	large
	Unrelated	1218 (104)	1384 (102)	-166 (46)	-3.60 (18)	0.002	0.848	large
100 ms	Related	1391 (108)	1555 (86)	-164 (43)	-3.74 (18)	0.001	0.882	large
	Unrelated	1214 (84)	1306 (78)	-91 (36)	-2.49 (18)	0.023	0.588	middle

Table. 21 The results of the independent test for the effect of proficiency on reaction times across all experimental conditions.

5.2.2 Accuracy analysis (ACC)

For the subject analysis, we conducted a mixed ANOVA analysis on the accuracy data in each condition with proficiency level (highly and less) and presentation times (50, 75, and 100 ms) as between-subjects factors and prime type (related and unrelated) as a within-subjects factor.

For the item analysis, we conducted a factorial ANOVA analysis on the accuracy data with prime type, presentation times, and proficiency level as within-subjects factors.

The ANOVA analysis showed a significant main effect for the type of the prime word in the analysis by subjects $F1(1, 54) = 13.75$, $MSE = 622.26$, $p < 0.001$, and in the analysis by items $F2(1, 29) = 14.61$, $MSE = 1711.74$, $p = 0.001$. Also, there was a significant main effect for prime presentation times in the analysis by subjects $F1(2, 54) = 13.42$, $MSE = 928.79$, $p < 0.001$, and in the analysis by items $F2\ (2, 58) = 17.55$, $MSE = 2611.88$, $p < 0.001$. Moreover, there was a marginal significant effect of proficiency level in the analysis by subjects $F1$ (1, 54) = 3.782, MSE = 261.71, p = 0.057, and a significant effect in the analysis by items $F2$ (1, 29) = 6.75, MSE = 795.07, p = 0.015. The interaction between the type of prime words and proficiency level was insignificant in the analysis by subjects $F1$(1, 54) = 0.53, MSE = 23.77, p = 0.472, and in the analysis by items $F2$ (1, 29) = 0.59, MSE = 95.07, p = 0.451. This suggests that the effects of prime type on accuracy did not differ essentially for either of the proficiency groups. Likewise, the interaction between presentation times and proficiency level was insignificant in the analysis by subjects $F1$ (2, 54) = 0.72, MSE = 49.69, p = 0.492, and in the analysis by items $F2$(2, 58) = 0.77, MSE =125.49, p = 0.468. This indicates that the effects of prime type on accuracy rate did not vary significantly for either of the proficiency groups. Finally, the interaction between presentation times and prime type was significant in the analysis by subjects F1(2,54) = 16.57, MSE = 749.71, p = 0.000, and in the analysis by items $F2$(2, 58) = 10.39, MSE = 2205.49, $p < 0.001$. This suggests that the effect of presentation times differs under the related and unrelated prime word conditions. As shown in figure 16, the accuracy rate under the unrelated condition dropped when the presentation times increased. However, for the related condition, the accuracy rate was high when presentation times were 50 ms and 75 ms, but lower when the presentation times increased to 100 ms. Finally, the three-way interaction between presentation times, prime type and proficiency level was insignificant in the analysis by subjects F1(2, 54) = 1.36, MSE = 61.34, p = 0.267, and in the analysis by items F2(2, 58) = 1.28, MSE = 179.65, p = 0.287. This suggests that the pattern of effect of the prime type and presentation times was the same for both proficiency groups.

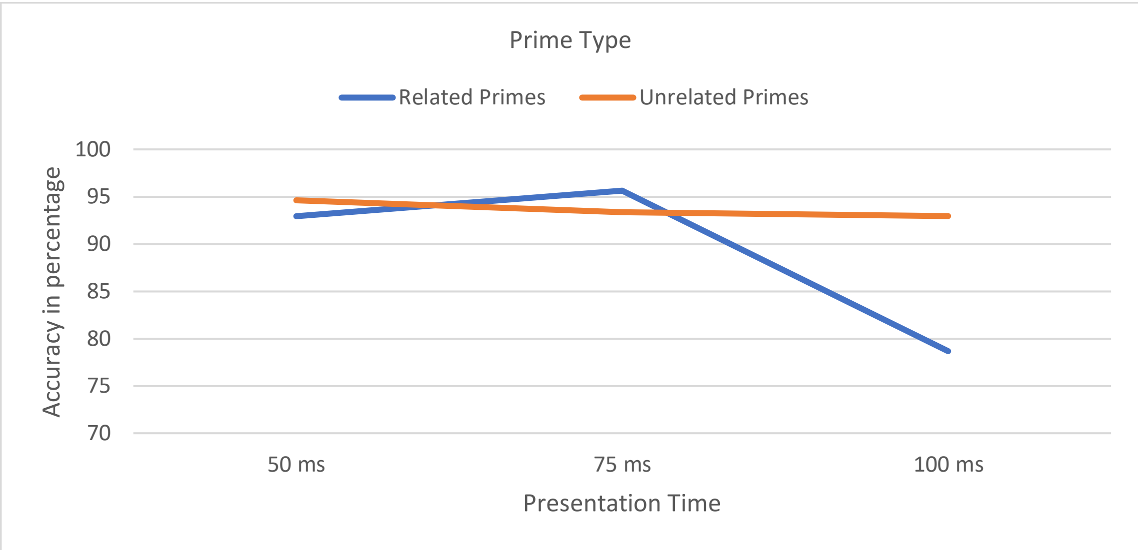


Figure. 16 The effects of presentation times on accuracy rate across related and unrelated prime type conditions.

We conducted a post-hoc test to examine the significance of the differences in accuracy rate between the several presentation times separately, under the related and unrelated prime words conditions (see Table 22).

Prime word type	Group 1	Group 2	Average ACC (SD)		Difference between group 1 & group 2 (SE)	T-test (df)	p-value (Bonferroni correction)	Effect size (Cohen's d)	Effect size tabulation
			Group 1	Group 2					
Related	50 ms	75 ms	92.93 (5.5)	95.64 (5.4)	-2.71 (1.73)	-1.57 (38)	0.249	0.255	small
	75 ms	100 ms	95.64 (5.4)	78.67 (9.7)	16.98 (2.50)	6.80 (38)	0.000	1.103	large
	50 ms	100 ms	92.93 (5.5)	78.67 (9.7)	14.26 (2.51)	5.69 (38)	0.000	0.924	large
Unrelated	50 ms	75 ms	94.62 (7.6)	93.33 (6.1)	1.29 (2.20)	0.59 (38)	1.124	0.095	small
	75 ms	100 ms	93.33 (6.1)	92.95 (10.02)	0.38 (2.63)	0.15 (38)	1.769	0.024	small
	50 ms	100 ms	94.62 (7.6)	92.95 (10.02)	1.67 (2.82)	0.59 (38)	> 0.999	0.096	small

Table. 22 The results of the independent sample t-test investigating the effect of presentation time on accuracy rate.

According to Table 22, the differences in accuracy rate between the three presentation times were not significant when the pictures were preceded by unrelated primes. In the related condition, the differences were insignificant when the pictures were presented for 50 ms and 75 ms; however, when the presentation time increased to 100 ms, the accuracy rate decreased significantly relative to the other presentation times.

Moreover, we conducted a paired t-test to assess the significance of the differences in accuracy rate between the related and unrelated distractors at each presentation time (Table 23).

Presentation times	Average ACC (SD)		Differences between the two groups (SE)	T-test (df)	P - value	Effect size (Cohen's d)	Effect size tabulation
	Related primes	Unrelated primes					
50 ms	92.93 (5.5)	94.62 (7.6)	-1.69 (2.16)	-0.78 (19)	0.443	0.180	small
75 ms	95.64 (5.4)	93.33 (6.1)	2.31 (1.39)	1.66 (19)	0.112	0.382	medium
100 ms	78.67 (9.7)	92.95 (10.02)	-14.28 (2.66)	-5.37 (19)	0.000	1.233	large

Table. 23 The results of the paired t-test to check the effect of the prime word type on accuracy rate across the different presentation times.

According to Table 23, the effect of the prime word type on accuracy rate was insignificant when the presentation time was 50 ms and 75 ms. This effect was significant only when the presentation time was 100 ms. In addition, we conducted an independent sample t-test to compare the performance of the two proficiency groups under the related and unrelated conditions (Table 24).

Prime word type	Average ACC (SD)		Difference between the two groups (SE)	T-test (df)	p-value	Effect size (Cohen's d)	Effect size tabulation
	Highly proficient	Less proficient					
Related	90.11 (9.4)	88.05 (11.19)	2.06 (2.67)	0.77 (58)	0.443	0.101	small
Unrelated	95.56 (6.1)	91.71 (9.2)	3.84 (2.02)	1.90 (58)	0.062	0.250	small

Table. 24 The results of the Independent sample t-test to compare the performance of the two proficiency groups in the related and unrelated conditions.

According to Table 24, the differences between the highly and less proficiency groups were statistically insignificant. This indicates that the effect of prime word type on accuracy was similar for both the highly and less proficient bilinguals. Likewise, the differences between the two groups were insignificant when evaluated across the various presentation times, as shown in Table 25.

Presentation Time	Prime word type	Average ACC (SD)		Difference between high & less (SE)	T-test (df)	p-value	Effect size (Cohen's d)	Effect size tabulation
		Highly proficient	Less proficient					
50 ms	Related	91.90 (6.8)	93.95 (3.7)	-2.05 (2.48)	-0.83 (18)	0.423	0.194	small
	Unrelated	97.33 (6.4)	91.90 (8.1)	5.43 (3.29)	1.65 (18)	0.117	0.389	small
75 ms	Related	96.62 (4.7)	94.67 (6.1)	1.95 (2.45)	0.80 (18)	0.436	0.188	small
	Unrelated	94.00 (6.6)	92.67 (5.8)	1.33 (2.79)	0.48 (18)	0.639	0.113	small
100 ms	Related	81.81 (9.4)	75.52 (9.4)	6.29 (4.23)	1.49 (18)	0.155	0.350	small
	Unrelated	95.33 (5.4)	90.56 (13.01)	4.77 (4.47)	1.07 (18)	0.300	0.252	small

Table. 25 The results of the independent t test to check the differences in accuracy between the two proficiency groups across the different presentation times and prime word types.

5.3 Discussion

The main objectives of this experiment were to first test whether semantic related prime words would induce interference in the masked primed picture naming task, and second to establish whether proficiency level would modulate the manner of lexical selection and cross language activation.

a) What is the manner of lexical activation/selection in different script bilinguals?

The results obtained in experiment two showed a significant main effect of prime words and presentation times, and a significant interaction between the type of the prime word and the presentation time. Longer naming latencies were detected when related prime words were presented for 100 ms and 75 ms before the target picture, indicating the presence of the semantic interference effect as typically found in the picture word interference task. The effect was not detectable when the presentation times were just 50 ms. However, the accuracy data shows a large effect of prime type at 100 ms presentation. In brief, longer exposure to related prime words induced inaccurate responses in both proficiency groups. The difference was also noticeable at 75 ms, but it did not reach a significant level.

Thus, these findings ruled out the argument that the cause of the semantic interference effect is a consequence of visually detected linguistic information coming from distractor words that force participants to form a verbal response to the distractor word prior to the target picture. We argue here that the longer naming latency is not only caused by the presence of the visual distractor, but also by the semantic relationship shared with the target picture. This result was consistent with studies that found a semantic interference effect when using the picture word interference tasks (Hermans et al. 1998; Hoshino 2006; Zhao et al. 2012), and the studies that used masked priming with monolingual participants (e.g., Alario et al. 2000; Bajo et al. 2003).

Next, we question what these results tell us about the nature of lexical access and selection by adult bilingual speakers? To answer this, we posit that finding semantic interference at 100 and 75 ms can be interpreted as evidence of lexical competition at the lemma level. To explain, we argue that the masked prime word initially activated the corresponding conceptual representations at the concept level, and then the activation spread to the corresponding lexical nodes at the lemma level, at which the competition occured between the target and non-target lexical nodes. We then asked why this effect was present at 100 ms and 75 ms, but not at the shorter presentation time (50 ms)? To answer this, we posit that when the masked prime was presented for 50 ms, the time available to process the prime word was only sufficient to partially activate the phonological representations, and thus did not activate the corresponding conceptual/lexical representation. By the time the conceptual representation of the target picture was fully active, any prior partial activation was ignored at the conceptual level.

Reflecting on the findings of this experiment, it is evident that the selection process takes into consideration both lexical nodes in the target and non-target language, and that there is a competition between lexical nodes at the lemma level. This contrasts with the findings from the first experiment demonstrating that the selection process does not consider non-target lexical nodes. We might explain this discrepancy as being a consequence of the different duration time (SOA). When testing the identity effect in experiment one, the prime words were presented for 50 ms and they yielded a facilitation effect; whereas when investigating the semantic interference effect in experiment two, the effect was found to occur at 100 ms but not at 50 ms. Thus, it could be argued that under longer durations: i.e., 100 and 75 ms, the non-target lexical nodes become highly activated, and this activation is high enough to impede the selection process, while under the short presentation times they fail to have an effect. However, this explanation may not hold, as Costa et al. (1999) reported an identity effect in the picture-word interference task when the translation distractor was presented for 200 ms at -200 SOA. Thus, further investigation is required to test the validity of the impact of duration of SOA.

Nevertheless, as pointed out in chapter four, there is an additional explanation for this discrepancy. The cause of this semantic interference effect might originate at the conceptual level. That is to suggest that the delayed naming is caused by competition at the conceptual level, due to the activation of several related concepts. In the next chapter (six), this explanation is investigated by comparing the performance of bilinguals in two tasks: one involving conceptual activation only, and the other conceptual and lexical activation. In these tasks, semantic relatedness was manipulated. If the effect originated at the conceptual level, then longer reaction times should be reported in both tasks. If the effect originates at the lexical level, then longer reaction times should be reported in the task involving lexical activation.

b) What is the effect of script differences on lexical access and manner of lexical selection/activation?

As discussed in section (a), the results of this experiment revealed that Arabic-English bilinguals experience a semantic interference effect, suggesting that different script bilinguals experience non-selective access, and that the non-target lexical nodes are sufficiently active at the lexical level to impede the selection process. This implies that the visually presented prime words boost the activation level of the non-target lexical nodes, which became high enough to compete for selection. Processing the prime words indicates that when a different script is perceptually available to bilinguals, it does not inhibit the cross-language activation. The finding concurs with that of the second experiment conducted by Boukadi et al. (2015), in which they tested Arabic- French bilinguals in a picture-word interference task, and reported a semantic interference effect. In Boukadi et al.'s (2015) experiment, the distractors were presented auditorily, i.e., they were not visually presented, yet they contributed to the cross-language activation. However, in a recent study, Hoshino et al. (2021) tested different script Japanese-English bilinguals in a picture-word interference task, in which the distractors were semantically related, or the translation equivalent. They found the presence of a translation facilitation effect (i.e., the identity effect) only, but no semantic interference effect. They explained that the target and non-target lexical nodes were active briefly, but then as the bilinguals exploited the language cues in the distractor words, the selection process became language specific. Thus, the bilinguals selected the target node earlier in their speech planning. We contribute the contradictory outcomes of the study by Hoshino et al. (2021) and the current experiment to differences in experimental procedures; the difference between our experiment and that of Hoshino et al. (2021) was the presentation modality of the distractor/prime words. In their study, Hoshino et al. (2021) presented the distractors word 25 ms after the presentation of the picture in red ink, and they remained on screen until the participants responded. In our study, the primes were masked and presented briefly for 100 ms before the presentation of the target picture, in order to produce the semantic effect. Thus, a longer visual exposure (unmasked) of different script distractors may facilitate lexical selection by inhibiting the activation of the non-target lexical nodes. However, further research is needed to validate this argument, perhaps by applying paradigms other than the picture-word interference task. For example, the effect of presenting masked vs unmasked related primes can be tested in a picture naming task in which the presentation time is manipulated (shorter vs longer exposure).

The findings of experiment two provide an important indication regarding the role of different script in lexical access and selection. The experiment to this point replicates the finding of the semantic interference effect in same script bilinguals, and provide evidence that non-target lexical nodes are active, regardless of script differences. In order to provide a clear picture of the effect of script differences on the manner of lexical activation and selection, further investigations were conducted in experiment three, four and five (chapters six, seven and eight, respectively).

c) What is the effect of proficiency level on lexical access and manner of lexical activation/selection?

Regarding the question of whether proficiency level modulates cross-language activation and manner of lexical selection**,** the results showed that priming the target pictures with related words did yield a semantic interference effect for both highly proficient and less proficient bilinguals. This indicates that proficiency level did not modulate cross-language activation nor the manner of lexical selection. An interesting finding here was that both groups experienced a semantic interference effect at 100 and 75 ms, but not at 50 ms, which means they both required the same processing time to show an effect. These findings contradict other studies (e.g., Kheder and Kaan 2019) claiming that proficiency modulates lexical selection, because highly proficient bilinguals apply a language specific selection whereas less proficient bilinguals apply language non-specific selection. However, proficiency level did influence retrieval timing for both groups, as the highly proficient participants were faster at naming the pictures relative to the less proficient bilinguals. Similar to experiment one, the results revealed a significant difference in the performance of the two groups, as the highly proficient bilinguals named the target pictures in all conditions more rapidly. Further, there was no significant interactions with other factors: i.e., prime type and presentation times. In addition, the analysis of accuracy data indicated that the highly proficient participants were more accurate. This leads us to question where the difference in the performance of the two groups with regard to naming latencies comes from?

We argue that there are two plausible explanations for these differences: the weak link as proposed by the Revised Hierarchical Model (RHM) (Kroll & Stewart, 1994; see also Kroll et al. 2010) and the Inhibitory Control mechanism proposed by Green (1998). The RHM model claims that greater experience with a language increases overall frequency of access over time; thus, the connection between the concept system and the L2 lexicon is weaker in less proficient bilinguals when compared with highly proficient bilinguals who access their L2 lexicon more frequently. Thus, the time required to retrieve L2 lexical nodes is longer for less proficient bilinguals compared to highly proficient bilinguals. The second explanation for this phenomenon is the Inhibitory Control (IC) mechanism proposed by Green (1998). According to language non-specific selection, the resolution of competition between two activated lexical nodes in the target and non-target language is achieved via a control mechanism that is responsible for controlling or suppressing the activation of the non-target lexical nodes at the lexical level. Additionally, it is assumed here that highly proficient bilinguals are consequently faster at resolving lexical competition due to their enhanced language control mechanism. Hence, it is logical to argue that the IC mechanism was operating, but that the strength of the suppression was dependent on the proficiency level of bilinguals: i.e., the highly proficient participants required less time to control the activation of non-target lexical nodes relative to the less proficient bilinguals.

When comparing the findings of experiments one and two, regarding the faster retrieval and accuracy of the highly proficient bilinguals compared with the less proficient bilinguals, the explanation proposed by the revised hierarchical model is compatible with the overall findings of the two experiments. The inhibitory control mechanism cannot account for the performance of the highly and less proficient bilinguals in experiment one, as the manner of selection was language specific. Thus, it seems that the Revised Hierarchical Model provides an accurate explanation for the pattern of findings in experiments one and two, regarding the better performance of the highly proficient bilinguals.

5.4 Conclusion

In summary, this study investigated whether a semantic interference effect would be apparent when testing Arabic-English bilinguals in a masked-priming picture naming task. Our purpose was to determine whether lexical selection in different script bilinguals entails competition, and whether proficiency level modulates lexical access and cross-language activation. The results showed that both groups experienced a semantic interference effect when naming pictures preceded by related primes. This indicated that both target and non-target lexical nodes were activated here and considered for selection. We argued that further investigation is essential to establish the root of this effect, whether it is located at the lexical level or at the conceptual level. Also, the findings indicated that proficiency level did not modulate cross-language activation among Arabic-English bilinguals. In the next chapter, we further expanded our investigation by combining two tasks (i.e., primed picture naming and animacy decision tasks), to locate the origin of the semantic interference effect.

Chapter 6 Experiment Three: Masked Priming of Non-Cognate pictures in a Naming Task and Animacy Decision Task

In experiments one and two, we investigated the manner of lexical selection during the production process, so as to determine whether the non-target lexical node competes with the target lexical node for selection or not. We used the masked priming method, and manipulated two experimental conditions (the identity effect and semantic relatedness). The findings from experiments one and two were contradictory as priming target words with translation words in L1 induced a facilitation effect; i.e., there was a lack of lexical competition between the target and non-target lexical nodes at the lexical level, whereas priming the targets with semantically related words in L1 induced an interference effect (i.e., created competition). We believe that these contradictory findings, are unlikely be a consequence of the sampling techniques or experimental procedures as they were carefully matched. This led us to consider the locus of the semantic interference/facilitation effect. We noted in the literature review (section 2.3.2), that several studies (e.g., Costa et al. 2005; Finkbeiner and Caramazza 2006; Mahon et al. 2007; Janssen et al. 2008) suggested that the interference effect found in picture-word interference tasks is located at the phonological level, and the facilitation effect at the conceptual level or lexical level. However, Abdel Rahman and Aristei (2010) found that semantic interference effect was present in tasks that do not involve phonological processing. This suggests the lexical level is a possible locus of the semantic interference effect. As discussed in section 2.3.2, we hypothesise that the conceptual level may be a better candidate for the locus of a semantic interference effect, because this is where semantic knowledge is stored. Thus, we asked firstly: What if semantic interference is caused by conceptual competition and not lexical competition? This is possible as different concepts are activated by the distractor word/prime word and the picture presented. Secondly, where does the semantic facilitation effect originate?

Recall, interference occurred when we primed the pictures with semantically related primes (cf chapter five), but facilitation occurred when we primed the pictures with their translation equivalents (cf chapter four). We question whether this is because the semantically related word has a different concept from the picture, whereas the translation equivalent shares the same concept, and thus instead of hindering the activation process it enhances it by sending additional activation to this concept thereby facilitating its selection?

To further explore this hypothesis, we contrasted these two alternative positions by exploring the semantic interference effect over two different tasks; (i) a task that requires access to conceptual, lexical and phonological level processing through to articulation, and (ii) a task that requires only conceptual processing.

In task one, the participants were presented with a series of pictures of an object (e.g., *apple*) preceded by an L1 masked prime word. They had to name these pictures in their L2. There were two experimental conditions: semantically related primes (e.g., *برتقال* [English orange]) in the critical condition and semantically unrelated primes (e.g., *ثعبان* [English snake]) in the unrelated condition. In this task, the conceptual, lexical and phonological features of the word are activated to a certain level based on its type (target vs non-target). As discussed previously, the interference effect observed in this task can be accounted for by either competition at the conceptual, lexical, or phonological level.

In the second task, the participants had to respond (by pressing a button yes or no) whether the identical pictures used in task one, were animate or inanimate objects. These pictures were preceded by semantically related or unrelated masked prime words in the L1. This animacy decision task involved conceptual activation only, as no lexical activation or overt naming of the pictures was required.

We hypothesized that if a semantic interference effect occurs at the conceptual level, there will be longer naming latencies in the picture naming task and in the animacy decision task when the pictures are preceded by semantically related words rather than unrelated words. However, if the semantic interference effect occurs at the lexical level, then there will be longer naming latencies across the semantically related condition in the picture naming task, but not in the animacy decision task. This is because the picture naming task involves lexical activation, unlike the animacy decision task which only involves the activation of concepts.

Similar to experiments one and two, in experiment three we examined the role of proficiency level and script differences on cross-language activation. We contrasted the performance of two groups of Arabic-English bilinguals (highly vs less proficient) across both tasks. In addition, we investigated whether bilinguals could exploit Arabic prime words as language cues and inhibit the activation of non-target language in the animacy decision task (which involves conceptual processing), and in the picture naming task (which involves conceptual, lexical and phonological processing).

6.1 Methods

6.1.1 Participants

The research participants were 182 adult volunteers, with eight participants excluded as their responses were not recorded. In addition, six more participants were removed owing to their high-test error rate with a very short reaction times (less than 100 ms), which indicated that the participants were not appropriately engaged with the task. So, in total, 168 participants were included in this experiment. The participants were all adult Arabic-English bilinguals. They were Saudi females for whom Arabic was their native language. All the participants had given consent to their data being used anonymously for research purposes prior to the study (see chapter three section 3.2), although none of the participants had contributed in experiment one or two.

In terms of the other experiments, the participants were divided in two groups according to their proficiency level (highly and less), as detailed in Chapter 3. An independent-samples t-test was conducted to compare the IELTS scores between the two groups. This resulted in the identification of a statistically significant difference between the highly proficient group (M = 7.5, SD = .515) and the less proficient group (M = 3.6, SD = .562), $t(166) = 46.387$, $p < .001$.

6.1.1.1 The results of the language history questionnaire

As summarised in Table 23, the participants' average age of acquisition was uniformly found to be around ten years old. The participants were also asked how they acquired English as a second language: whether through formal classroom teaching and/or interactions with other people. The results revealed that the highly proficient participants had primarily acquired English as a second language through a mixture of formal classroom teaching and interactions with other people (mostly native English speakers). This was because they had acquired their postgraduate degree from universities in: the United Kingdom, the United States of America, and/or Australia, meaning they had lived in an English dominant environment for a minimum of two years. On the other hand, the less proficient participants reported that they acquired English as a second language principally through classroom teaching. In addition, they had received instruction in Arabic throughout their early school education, i.e., in their elementary and intermediate school.

Moreover, the participants were asked to estimate their daily use of L1 and L2. As shown in Table 26, the daily average estimation of the use of L1 and L2 by the less proficient group was 71.22% and 28.77%, respectively, whereas the average estimation for the highly proficient group for L1 and L2 usage per day was 29.87% and 70.12%, respectively. Table (26) shows the descriptive statistics for each question across the two groups. The data extracted from the language history questionnaire showed that the groups differed in terms of their daily use of their L1 and L2, language of instruction at high school and university, their age and their level of education.

However, the age at which they started learning English was similar. In summary, the data indicated that the highly proficient group had been and continue to be more exposed to English as a second language than the less proficient group.

The Highly Proficient Group (n=82)	**Minimum**	**Maximum**	**Mean**	**Std.Deviation**
Age (years)	26	43	33.18	5.06
IELTS	6.5	8.5	7.15	0.569
Mean daily L1 usage (%) (5 pt scale)	>25	50	29.8	15.4
Mean daily L2 usage (%) (5 pt scale)	50	100	70.122	15.43
Age of acquisition (years)	3	10	10.488	2.3425

The Less Proficient Group (n=86)	**Minimum**	**Maximum**	**Mean**	**Std.Deviation**
Age (years)	19	20	19.5	0.826
IELTS	3	4.5	3.64	0.562
Mean daily L1 usage (%) (5 pt scale)	50	75	71.28	9
Mean daily L2 usage (%) (5 pt scale)	25	50	28.77	9.01
Age of acquisition (years)	3	13	10.279	2.2786

Table. 26 Descriptive statistics for participants in both groups.

6.1.1.2 The results for the self-rating

The average scores for both groups for each skill were shown in table 27. An independent-samples t-test was conducted to compare the mean scores for the self-assessed rating in each skill for the two groups, leading to the identification of statistically significant differences between the two groups across all the skills: (i) Reading $t(164.4) = 30.25$, $p < .001$, (ii) writing $t(153.4) = 22.36$, $p < .001$, (iii) speaking $t(166) = 26.59$, $p < .001$, and listening $t(166) = 23.93$, $p < .001$. The results suggested that, based on their own perception of their proficiency level, the highly proficient bilinguals considered themselves native like; whereas the less proficient bilinguals considered themselves to be maintaining a basic level of proficiency.

Group Statistics	Proficiency	N	Mean	Std. Deviation	Std. Error Mean
Self-Rate Reading (5 pt scale)	High	82	4.59	0.496	0.055
	Less	86	2.33	0.471	0.051
Self-Rate Writing (5 pt scale)	High	82	4.26	0.492	0.054
	Less	86	2.19	0.695	0.075
Self-Rate Speaking (5 pt scale)	High	82	4.3	0.463	0.051
	Less	86	1.91	0.68	0.073
Self-Rate Listening (5 pt scale)	High	82	3.96	0.508	0.056
	Less	86	2	0.553	0.06

Table. 27 Descriptive statistics for the self-assessed rating in participants' skills in L2 across the two groups.

6.1.1.3 The results of the lexical decision task

An independent-group t-test was performed on latencies and accuracy for L2 words and nonwords with proficiency (highly proficient bilinguals vs. less proficient bilinguals) as an independent variable. The highly proficient bilinguals were faster and more accurate for nonwords than the less proficient bilinguals [$t(120.4) = 14.68$, $p < .001$ for latencies; $t(166) =$

16.41, p < .001 for accuracy]. In addition, the highly proficient bilinguals were significantly faster and more accurate for words than the less proficient bilinguals [$t(135.68) = 10.978$, p < .001 for latencies; $t(109) = 13.24$, p < .001 for accuracy]. These results suggested that there was a significant difference between the two groups in terms of L2 proficiency. The mean accuracy and reaction times for words and non-words are shown in Table 28.

	Non-word		**Word**	
Proficiency Groups	Reaction Times	Accuracy in percentage	Reaction	Accuracy in percentage
Highly Proficient Bilinguals (n=82)	921 (166)	95.27 (9.4)	769 (161)	98.04 (3.12)
Less Proficient Bilinguals (n=86)	1566 (365)	72.97 (8.2)	1156 (285)	85.29 (8.2)

Table. 28 Lexical decision results for the highly and less proficient bilinguals in experiment three. Standard deviations are in parenthesis.

To conclude, the results of the language history questionnaire, the self-assessed rating and the lexical decision task suggest that the two groups varied significantly in terms of their proficiency level.

6.1.2 Materials

6.1.2.1 The selection of the pictures

Sixty-six stimuli words were adopted from the Prototypically Norms for 26 Semantic Categories (Uyeda and Mandler 1980). These words were used to construct: (i) a target stimuli list with forty words; (ii) a practice list with six words; and (iii) a filler item list with twenty words. The words were the names of objects from the different semantic categories (e.g., human, animals, furniture, food, clothing, and musical instruments) and all were non-cognates. Half the total word list (i.e., thirty-three) named animate objects, and the other half inanimate objects. Then, sixty-six corresponding black and white line-drawings were selected from the International Picture Naming Project online-database (Szekely et al. 2003; Szekely et al. 2004) (see appendix K for pictures samples). Each animate picture name was carefully matched with an inanimate picture name on variables that are often considered critical to the type of processing level involved in the task and the experimental conditions (for details, see section 3.5.1). In brief, the picture names were controlled at *p*s > .05(see Table 29 for these means).

Variables	**Animate list (n = 20)**	**Inanimate List (n = 20)**	***t-test***	***p* value (t-test)**
Name Agreement (%)	0.93 (.079)	0.93 (.122)	*t*(38) = .122	*p* > .05
Visual Complexity (KB)	20235.15 (9114.5)	19192.85 (7956.2)	*t*(38) = .385	*p* > .05
Syllable Length	1.50 (.688)	1.90 (.968)	*t*(38) = -1.506	*p* > .05
CharacterLength	5.00 (1.686)	6.30 (2.736)	*t*(38) = -1.81	*p* > .05
Frequency per million words	3.35 (1.13)	3.21 (1.40)	*t*(38) = .346	*p* > .05
Age of Acquisition (1-3 point scale)	1.75 (.967)	1.90 (.968)	*t*(38) = -.490	*p* > .05
Imageability (100-700)	528.26 (52.5)	552.68 (54.5)	*t*(38) = -1.41	*p* > .05
Familiarity (100-700)	615.37 (16.30)	600.39 (35.16)	*t*(38) = 1.67	*p* > .05

Table. 29 Characteristics of pictures and picture names in English used in experiment three.

6.1.2.2 The selection of the prime words

With regard to the selection of the prime words, for each target stimuli, we chose: (i) a related word from the same semantic category to form a related prime; and (ii) another unrelated word from a different semantic category to form an unrelated prime. In addition, six unrelated prime words were selected for the practice list, and an additional twenty unrelated prime words were selected for the filler items, each adopted from the same database (for a full list, see appendix G).

The chosen prime words were translated into Arabic. We ensured that all the Arabic prime words and target picture names had a different onset and did not rhyme. Moreover, each related prime word was carefully matched with an unrelated prime word on number of letters, number of syllables, and frequency (see section 3.5.2). Related and unrelated prime word lists were similar in their frequency, number of syllables, and number of characters (see Table 30 for these means) (all ps > .05).

Variables	**Related Primes M (SD)**	**Unrelated Primes M (SD)**	***p* value (t-test)**
Syllable Length	2.28 (.857)	2 (.805)	$p > .05$
character Length	5.44 (1.46)	5 (1.469)	$p > .05$
Written Form Frequency	10.62 (30.20)	19.53 (33.35)	$p > .05$

Table. 30 Characteristics of L1 prime words in used in experiment three.

To assess the semantic relatedness of the chosen prime words to the target picture names, we asked a group of seventy-eight Arabic-English speakers to judge the similarity of the prime-target picture name pairs in terms of meaning on a 5-point Likert scale; where “1” meant very different and “5” very similar. Each participant was presented with eighty prime-target picture name pairs: forty were prime-target picture pairs for the semantically related condition, and forty prime-target picture pairs for the unrelated condition. The presentation of these prime-target name pairs was randomized.

The mean rating for each prime-target word pair was calculated and analysed. The results from the paired samples t-tests showed a significant difference in the rating scores, [$t(39) = 48.076$, $p < 0.001$]. The semantically related list was rated as more similar than the unrelated list (see Table 31).

Conditions	**Mean**	**SD**
Semantically Related Pairs (n 40)	4.22	.378
Unrelated Pairs (n 40)	1.38	.224

Table. 31 The similarity rating for prime-target pair

6.1.2.3 The Organization of Material for Animacy decision and picture naming task

The final stage of the organization process involved counterbalancing the presentation of the stimuli list across the participants and controlling for the animacy factor over the two tasks. For the animacy decision task, we kept the list of stimuli divided in half, according to their animacy status. When organizing the prime words in the related condition, we assigned animate-related primes to animate target pictures and inanimate-related primes to inanimate target pictures. Whereas, in the unrelated condition, we assigned a combination of animate/inanimate unrelated primes to the target pictures, regardless of their animacy status. Thus, an animate picture could be preceded by an animate/inanimate unrelated prime. If unrelated prime words were always matched to the target pictures in terms of their animacy, this would affect their decision, as the participants could use primes as cues to facilitate their decision. We created two versions of stimuli A and B to counterbalance their presentations across the participants. Therefore, the pictures names primed with related primes in version A were primed with unrelated primes in version B, and those primed with unrelated primes in A were primed with the related primes in version B. The participants completing the animacy decision task were then divided into two groups. The first group was assigned to version A, and the second group to version B.

For the picture naming task, we used the identical stimuli-prime pairs to those used in the animacy decision task, but we organized the list differently to fit the task. More specifically, we created two lists: each list with 20 animate and 20 inanimate pictures names. Prime-target name pairs were organized as follow:

(i) The semantically related condition consisted of 10 animate and 10 inanimate pictures, preceded by a semantically related prime;

(ii) The unrelated condition consisted of 10 animate and 10 inanimate pictures were preceded by unrelated primes.

The pictures preceded by related primes in the first list were preceded with unrelated primes in the second list, and those preceded by unrelated primes in the first list were preceded by a related prime in the second list. The prime-target picture pairs were counterbalanced across the group of participants. The participants were also divided into two groups, whereby the pictures assigned for the related condition in this group were assigned later for the unrelated condition in the second group.

6.1.3 Designs

The experiment was a mixed design (i.e., a 2 x 2 x 2 factorial design) with proficiency level (highly proficient bilinguals and less proficient bilinguals) and the type of the task (i.e., picture naming and animacy decision task) representing the between-subjects factors and prime type representing the within-subjects factor (related and unrelated primes). Thus, half the participants in both the less proficient group and the highly proficient group were asked to name pictures and the other half were asked to respond regarding whether the picture was animate or non-animate in the animacy decision task.

6.1.4 Experimental procedures for picture naming and animacy decision tasks:

This experiment is administered online unlike experiment one and two (see for details chapter three, section 3.4.2).

a) Masked priming in a picture naming task

We used the same procedures here as applied in experiment one and two. The only change we made was that an audio tone was presented at the onset of the target picture presentation. As explained in chapter three (section 3.4.2), the tone was added to help measure the participants' reaction times using Praat software.

b) Animacy decision Task

Similar to the picture naming task, pictures and words were shown in black on a white background. The picture size was 300 x 300 cm, and the font of the word was 48 pts. The participants were informed that they were required to look at the centre of the screen to view the picture presented. They were instructed to press the yes button on the screen if the picture represented something animate (i.e., a human being or animal), and the no button if the picture represented something inanimate (i.e., a musical instrument, furniture, etc.). They were not informed of the presence of the primes, which occurred prior to the presentation of the pictures during the task. Example pictures of living and non-living things were provided during the oral instruction phase. Then the participants completed a six- questions training session, which was performed with separate stimuli and prime words that differed from those employed in the actual test. Each trial consisted of the following sequence of events:

(i) The tests initially commenced with a fixation point (+), which appeared in the middle of the screen for 500 ms;

(ii) A visual mask of (######) symbols then replaced the fixation point, and remained for 500 ms;

(iii) A prime word (related or unrelated) appeared for 100 ms;

(iv) A visual mask of ########) symbols remained for 14 ms;

(v) A target picture which remained on screen until the participants responded by clicking the yes/no button on the screen (mouse click or touch click).

6.2 Analysis of Results

The data were collected from one hundred and sixty-eight participants. After trimming the data (described in section 3.6.1), we calculated the reaction times, and accuracy for correct responses again. Errors at 1.9 % and outliers at 2.2 %, in the picture naming task and errors at 2.1 %, and outliers at 1.0% for the animacy decision task were excluded from the following analyses. The mean reaction times and percentage of accuracy are shown in Table 32.

Task	Prime Type	Highly Proficient		Less Proficient	
		Reaction times	accuracy in percentage	Reaction times	accuracy in percentage
Picture naming	Semantically related	1872 (195)	99.13 (3.8)	1963 (273)	97.07 (4.4)
	Semantically unrelated	1125 (312)	98.97 (4.0)	1510 (275)	96.95 (4.7)
Animacy decision	Semantically related	1040 (263)	98.46 (4.4)	1026 (253)	96.93 (6.8)
	Semantically unrelated	1777 (421)	99.24 (2.1)	1159 (196)	97.04 (5.02)

Table. 32 Reaction times and accuracy data for the highly and less proficient bilinguals across in the two experimental conditions across the tasks. Standard deviations are in parentheses.

6.2.1 Reaction times analysis (RT)

A mixed ANOVA analysis was performed on the mean response latencies per subject with (a) proficiency level (highly proficient vs less proficient) and task type (animacy decision task vs picture naming task) as the between-subjects factors, and (b) prime type (semantically related and unrelated prime words) as the within-subjects factors. In addition, a repeated ANOVA was performed on the mean response latencies per items in each condition.

The results showed a significant main effect of prime type in the analysis by subjects $F1$ (1, 164) = 8.57, MSE = 0.57, p = 0.004, and in the analysis by items $F2$ (1, 152) = 47.51, MSE = 0.56, p < 0.001. A significant interaction was found between the type of prime word, proficiency level and the task type in the analysis by subjects $F1$ (1, 164) = 63.24, MSE = 4.23, p < 0.001, and in the analysis by items $F2$ (1, 152) = 328.86, MSE = 3.89, p < 0.001. Therefore, reporting the main effects is misleading, because the impact of the prime word could vary (and even become contradictory) in different subgroups.

Also, a significant interaction was found between prime type and task type in the analysis by subjects $F1(1, 164) = 335.16$, MSE = 22.41, $p<0.001$, and in the analysis by items $F2(1, 152) = 45.09$, MSE = 0.53, $p < 0.001$. This interaction suggests that the effect of the prime words differed across tasks, and was dependent on the nature of the task. Figures 17 and 18 show the different effects of prime type on reaction times across the different tasks and proficiency levels. More specifically, in the animacy decision task, the related prime words induced faster naming latencies for the highly proficient group and the less proficient group relative to the unrelated prime words (Figure 17). Reaction times were considerably longer in the case of unrelated prime words for the highly proficient group when contrasted with the less proficient group. In other words, the size of the semantic effect is greater for the highly proficient group than the less proficient group and we discuss the implication of this finding in section 6.3.

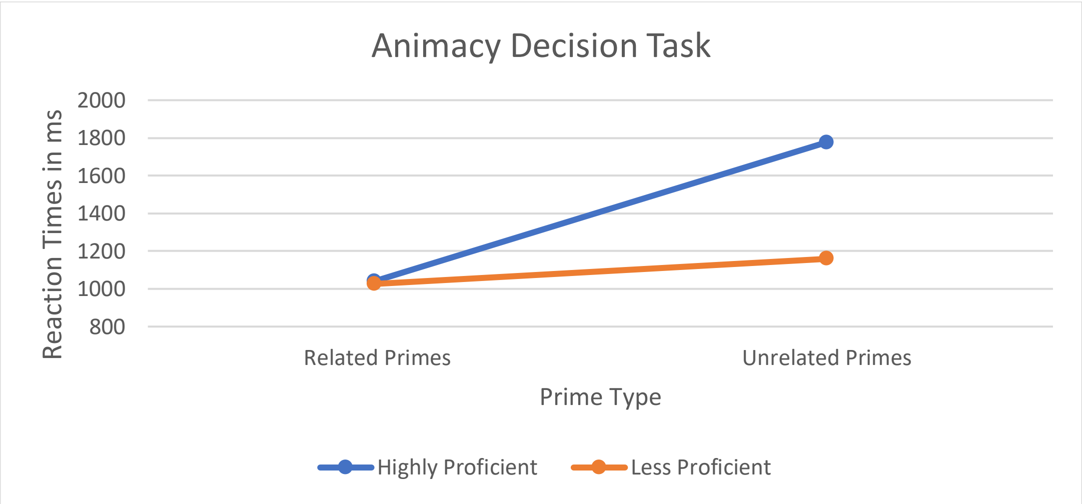


Figure. 17 The effect of prime type on reaction times for different proficiency groups in cases of the animacy judgment task.

A different pattern was observed for the picture naming task. As shown in Figure 18, related prime words produced longer naming latencies for both the highly proficient and less proficient bilinguals. Similarly, reaction times were shorter for the unrelated prime words when compared with related prime words for both proficiency groups. Similar to the animacy decision task, the size of the semantic effect was greater for the highly proficiency group when compared to the less proficient group.

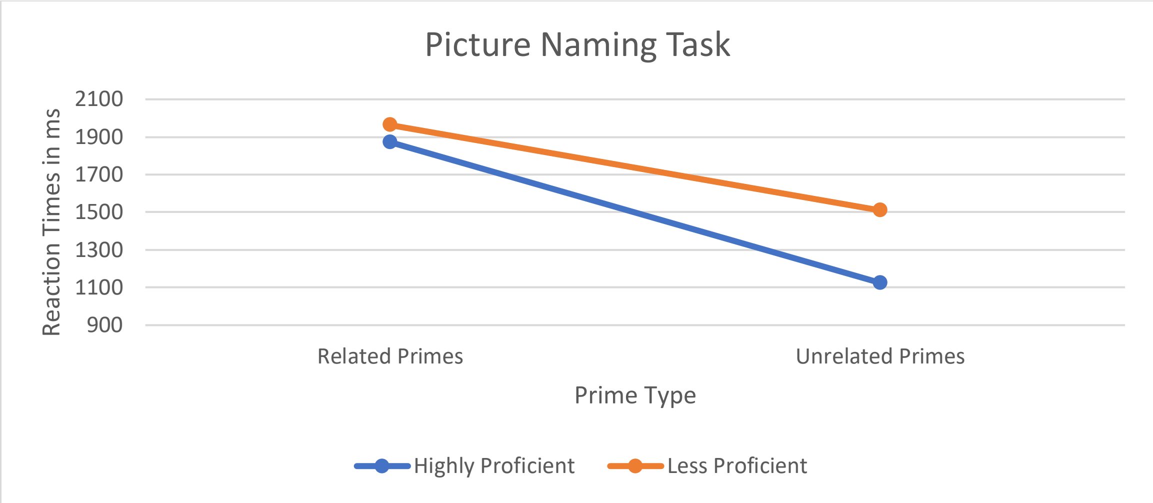


Figure. 18 The effect of Stimulus on RT for different Proficiency groups in cases of Picture tasks.

In addition, the partial eta-squared values were close to 0.06 for the effects of prime type, and the interaction between prime type and proficiency level, indicating a "small" effect size in Cohen's terms (1988). The values of eta-squared were > 0.14 for the differences in the effect of prime type for different proficiency-task subgroups, indicating a large effect size.

We conducted a paired t-test to examine the significance of the differences noted for various tasks and proficiency levels. Bootstrapping was used (1000 samples) to avoid any issues with data distribution (cf.,Field 2013). As summarised in Table 33, the differences were statistically significant at the level 0.05. The calculated Cohen's d was > 0.8, indicating a large effect size for prime type in all cases excluding the less proficient subgroup in the animacy naming task where the Cohen's d was around 0.5, indicating a medium effect size. The implications of this are discussed in section 6.3.

Proficiency Level	Task Type	Average RT (SD) Related	Average RT (SD) Unrelated	Difference between related & unrelated (SE)	T-test (df)	p-value of t-test	Effect size (Cohen's d)	Effect size tabulation
Highly proficient	Animacy decision	1040 (263)	1777 (421)	-737 (83)	8.810 (39)	0.001	3.209	large
	Picture naming	1872 (195)	1125 (312)	747 (53)	14.14 (39)	0.001	3.468	large
Less proficient	Animacy decision	1026 (253)	1159 (196)	-133 (20)	-6.341 (45)	0.001	0.522	medium
	Picture naming	1963 (273)	1510 (275)	453 (57)	7.825 (41)	0.001	3.309	large

Table. 33 Paired t-test results to test the effect of semantically related primes on RT across different proficiency-task subgroups.

As mentioned earlier, there was a significant interaction between proficiency level and task type as a between-subjects effect indicating that variations in reaction times between highly and less proficient bilinguals are a consequence of differences in task type. The corresponding eta-squared has a value of 0.309 (> 0.14), indicating a large effect size describing these differences. Figure 17 shows that reaction times are approximately equal for both groups (highly and less) within the task "Animacy decision" in the case of related primes, and that the reaction times for the highly proficient group is longer than the reaction times for the less proficient group in the case of unrelated primes. Figure 18 indicates that the less proficient bilinguals had longer reaction times than the highly proficient bilinguals when completing the picture naming task in both conditions, regarding related and unrelated primes.

We conducted a paired follow-up independent sample t-test with bootstrapping to test the significance of the above-noted differences (Table 34). As shown in Table 34, the difference between the two groups of bilinguals was insignificant for both tasks when the primes were semantically related. The less proficient bilinguals demonstrated longer reaction times compared with the highly proficient bilinguals in both the picture naming task and the semantically related condition, but shorter reaction times in the animacy decision task in the case of unrelated stimulus.

Task Type	Prime Type	Average RT (SD)		Difference between highly & less (SE)	T-test (df)	P - value of t-test	Effect size (Cohen's d)	Effect size tabulation
		Highly proficient	Less proficient					
Animacy decision	Related	1040 (263)	1026 (235)	14.5 (53)	.271 (84)	0.799	0.058	micro (< small)
	Unrelated	1777 (421)	1159 (196)	618 (69)	8.507 (53.43)	0.001	2.001	large
Picture naming	Related	1872 (195)	1963 (273)	-90 (52)	-1.727 (74.3)	0.108	0.385	small
	Unrelated	1125 (312)	1510 (275)	-385 (65)	-5.926 (80)	0.001	1.310	large

Table. 34 Independent samples t-test results to test differences in RT between high and less proficiency groups across different tasks and Stimulus

6.2.2 Accuracy Analysis (ACC)

The results of mixed ANOVA indicated no significant impact of prime type on accuracy in the analysis by subjects $F1(1, 164) = 0.12$, $MSE = 1.88$, $p = 0.735$, and in the analysis by items $F2(1, 152) = 0.20$, $MSE = 2.27$, $p = 0.654$. Moreover, there was no significant interactions between prime type and proficiency level in the analysis by subjects $F1(1, 164) = 0.13$, $MSE = 2.11$, $p = 0.720$, and in the analysis by items $F2(1, 152) = 0.07$, $MSE = 0.77$, $p = 0.795$. Similarly, there was no significant interaction between prime type and task type in the analysis by subjects $F1\ (1, 164) = 0.43$, $MSE = 7.06$, $p = 0.512$, and in the analysis by items $F2(1, 152)$

= 0.36, *MSE* = 4.06, *p* = 0.550. Finally, no interaction was found between prime type, proficiency level and task type in the analysis by subjects *F1*(1, 164) = 0.16, *MSE* = 2.64, *p* = 0.688, and in the analysis by items *F*2(1, 152) = 0.59, *MSE* = 6.63, *p* = 0.445.

The follow up paired t-test indicated no significant impact of prime type on accuracy across different tasks and proficiency levels (Table 35).

Proficiency Level	Task Type	Average RT (SD)		Difference between related & unrelated (SE)	T-test (df)	p-value	Effect size (Cohen's d)	Effect size tabulation
		Related	Unrelated					
Highly Proficient	Animacy decision	98.47 (4.4)	99.24 (2.1)	-0.78 (0.70)	-1.03 (39)	0.330	0.187	small
	Picture naming	99.13 (3.8)	98.97 (4.0)	0.16 (0.44)	.336 (39)	0.741	0.030	< small (micro)
Less Proficient	Animacy decision	96.94 (6.8)	97.04 (5.02)	-0.10 (1.15)	-.090 (45)	0.937	0.018	< small (micro)
	Picture naming	97.08 (4.4)	96.96 (04.7)	0.12 (0.87)	.136 (41)	0.887	0.055	micro

Table. 35 Paired t-test results to test the effect of prime type on accuracy across different proficiency-task subgroups.

In addition, the results of ANOVA indicated a significant main effect of proficiency level in the analysis by subjects *F1*(1, 164) =11.64, *p* = 0.001, and in the analysis by items *F*2(1, 152) = 0.59, *MSE* = 6.63, *p* = 0.445. There was no significant main effect of task type or significant interaction between proficiency level and task type (*p* =0.884 and *p* = 0.882 respectively). As shown in Table 36, we observed a higher accuracy rate for the more proficient group across the different tasks and conditions, and these differences were significant for the animacy decision task in the case of unrelated primes, and in the picture naming task in the semantically related prime condition. The effect size is classified as medium in Cohen's terms, as Cohen's d values were around 0.5 (Cohen 1988).

Task Type	Prime Type	Average RT (SD)		Difference between High & Less (SE)	T-test (df)	p-val of t-test	Effect size (Cohen's d)	Effect size tabulation
		High proficiency	Less proficiency					
Animacy decision	Related	98.47 (4.4)	96.94 (6.8)	1.53 (1.23)	1.20 (84)	0.227	0.270	small
	Unrelated	99.24 (2.1)	97.04 (5.02)	2.20 (0.79)	2.5 (84)	0.016	0.614	medium
Picture naming	Related	99.13 (3.8)	97.08 (4.4)	2.05 (0.91)	2.2 (80)	0.030	0.494	medium
	Unrelated	98.97 (4.0)	96.96 (04.7)	2.02 (0.98)	2.08 (80)	0.054	0.461	medium

Table. 36 Independent samples t-test results to test differences in RT between high and less proficiency groups across different tasks and prime type

6.3 Discussion

a) What is the locus of the semantic interference effect?

To determine the locus of the semantic effect (interference and facilitation effect) for word production in bilinguals, we tested the effect of semantically related primes on the performance of two groups of Arabic English bilinguals (highly proficient and less proficient) across two different tasks, namely the picture naming task and the animacy decision task. We hypothesized that if the semantic interference effect originates at the conceptual level, then longer naming latencies should be evident in the picture naming task and in the animacy decision task when pictures are preceded by semantically related words that are relevant to unrelated words. However, if the semantic interference effect originates at the lexical level, then longer naming latencies should be notable in the semantically related condition in the picture naming task only, but not in the animacy decision task. This is because the picture naming task involves lexical activation, unlike the animacy decision task which involves only the activation of concepts.

The results of experiment three revealed a significant main effect of prime type, and a significant interaction between prime type and task type. This means that the type of prime words considerably affected the reaction times in both tasks, and the nature of the effect varied across the two tasks. In the masked primed picture naming task, semantically related prime words induced longer naming latencies relative to the semantically unrelated prime words. Whereas, in the animacy decision task, the semantically related primes induced a shorter reaction times relative to the semantically unrelated primes. The same pattern of effect was present for both highly proficient and less proficient bilinguals. The only distinction was that the size of the semantic effect was greater for the highly proficient bilinguals in both tasks.

The findings suggest that semantically related primes induced interference in the picture naming task, whereas in the animacy decision task, they induced facilitation. This pattern of findings can be interpreted as evidence supporting the view that the locus of the semantic interference effect is at the lexical level, because the effect was evident only in the task that required lexical activation. Thus, we argue that the semantically related prime word activated its concept at the conceptual level from among other activated concepts (the concept of the target picture and other related concepts). These semantically active related concepts send activation to the target concept at the conceptual level, and to their corresponding lexical nodes at the lexical level too. Thus, at the conceptual level, several concepts are active, which, according to the parallel activation level, will activate all the corresponding lexical nodes. Thus, the activated lexical nodes, at the lexical level, will compete for selection. For example, if a target picture of a *strawberry* is primed with a related prime word in L1 (e.g., توت [English cranberry]), the target concept, related concepts and the prime word concept will be activated at the conceptual level (Figure 19). The target picture will activate its corresponding concept and other related concepts e.g., *(‘grapes’, ‘blueberries’*, etc). The prime word *cranberry* will send activation to its corresponding concept at the conceptual level. All the activated concepts will then send activation to the corresponding lexical nodes in the target language and the non-target language. Now, there are several active lexical nodes in the target and non-target language at the lexical level, which in turn will compete for selection.

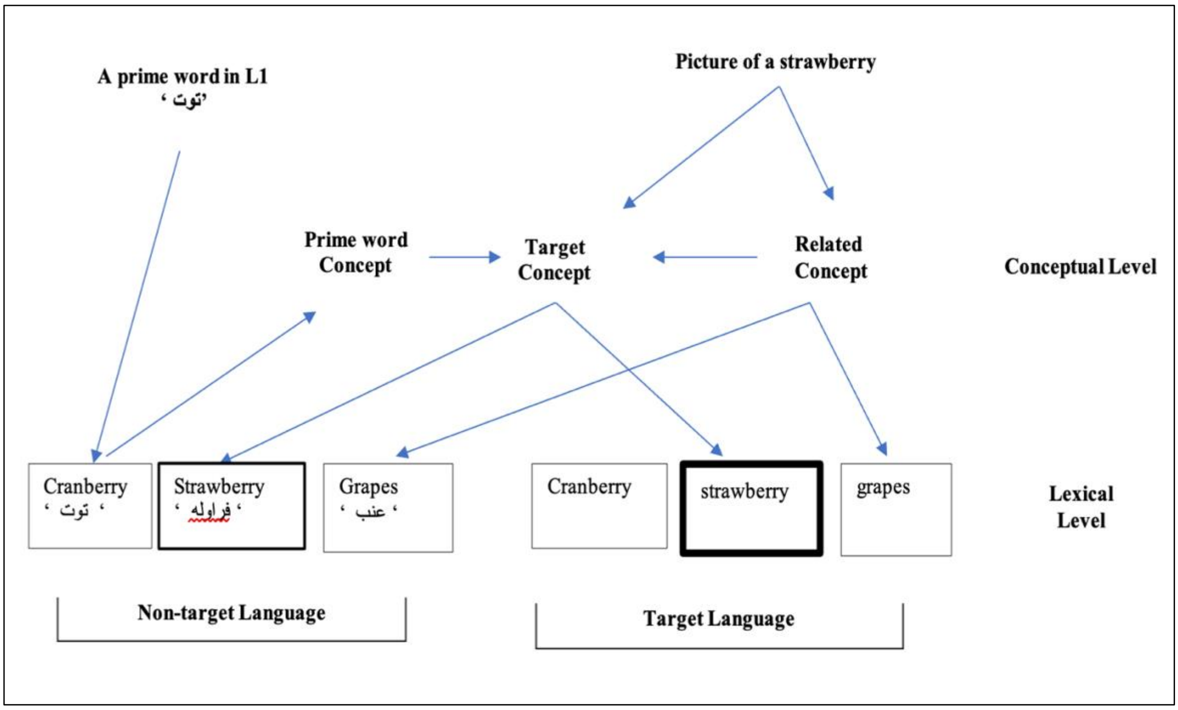


Figure. 19 A schematic representation of picture naming in L2 primed with semantically related masked words in L1. The arrows represent the flow of activation, while the thickness of the shapes represents the level of activation of the representations.

The other significant finding concerning the semantic facilitation effect in the animacy decision task supported our explanation that the interference effect manifests only if the task requires lexical activation. In the animacy decision task, semantically related concepts were not required to activate corresponding lexical nodes, as the task was purely conceptual. A possible explanation of this finding is that the activated concepts send activation to the target concept, which will eventually lead to facilitation not interference (Figure 20), since no activation of the target lexical representation is required here. Thus, the pattern of findings suggests that activating semantically related concepts will result in interference if the task requires the lexical retrieval of a target node or processing, whereas, if it does not, which is the case in an animacy decision task, it will lead to facilitation only.

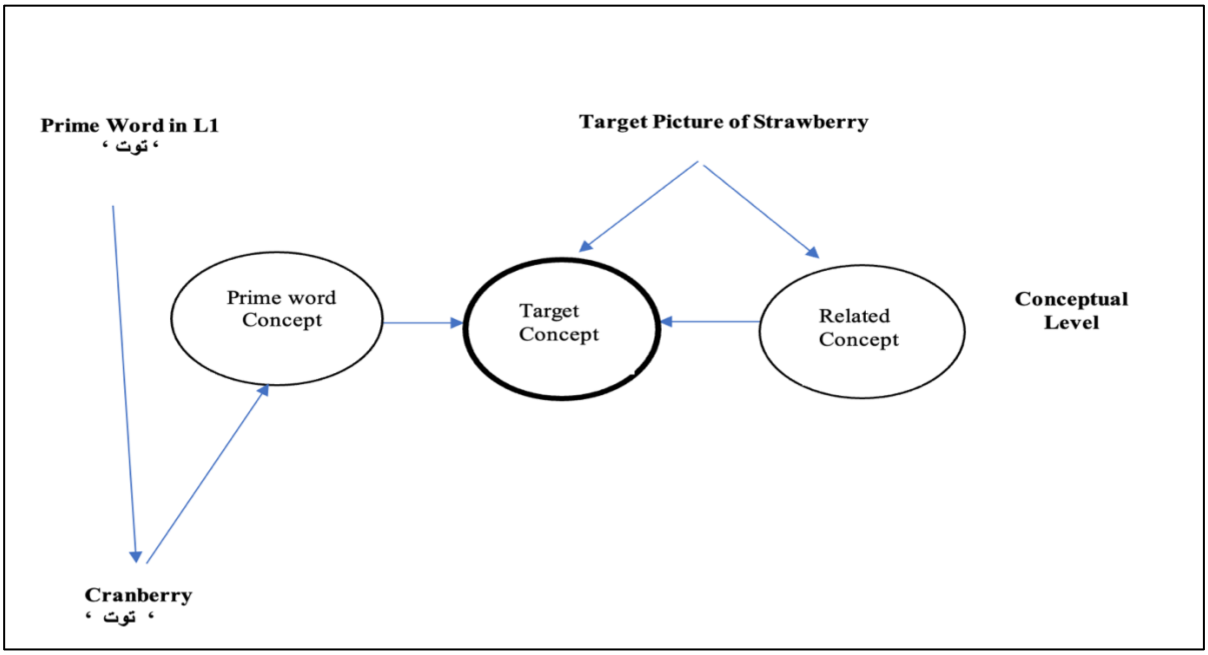


Figure. 20 A schematic representation of Animacy Decision task primed with semantically related masked words in L1. The arrows represent the flow of activation, while the thickness of the shapes represents the level of activation of the representations.

Considering the accuracy data, there was an impact on accuracy rate in all experimental conditions in both tasks: animacy decision and picture naming; however this impact failed to reach statistical significance. This means that the semantically related primes in this experiment did not affect the accuracy rate of participants' responses. However, the lack of a significant impact on accuracy data does not negate the overall effect of semantic manipulation on participants' performance. Notably, longer naming latencies are valid evidence for the existence of an effect.

The findings of this study are in line with the work of Bajo et al. (2003), who conducted research into the origin of the semantic interference effect. They tested the resulting effect when presenting a semantically related prime in a gender decision task, which involves processing at the lemma level. The performance of the Spanish monolinguals was affected by the presence of the related prime which suggests that the source of this effect is at the lemma level (i.e., lexical level).

b) What is the effect of proficiency level on lexical access and manner of lexical selection/activation?

When comparing the performance of both groups of bilinguals, no significant difference was noted in their reaction times overall. However, a significant interaction was found between proficiency level, prime type and task type. In the animacy decision task, the highly proficient bilinguals responded as quickly as the less proficient bilinguals in the related condition. However, the less proficient bilinguals outperformed the highly proficient bilinguals in the unrelated condition. In other words, the responses of the highly proficient bilinguals were significantly slower than the less proficient bilinguals in the unrelated condition. This is because in the unrelated condition, we assigned a mixture of animate/inanimate unrelated primes to the target pictures, regardless of their animacy status. It appears that the highly proficient bilinguals were able to fully access the semantic features of the prime words. Since the task required retrieval of the animacy status of the target picture, it is possible that they retrieved the animacy status of the prime words as well. Having different animacy categories (prime-target pair) resulted in the observed delayed responses, as the highly proficient participants had to ignore/supress the non-target activated concept. Thus, the effect size of semantic facilitation was greater for the highly proficient bilinguals than it was for the less proficient bilinguals (737 ms and 133 ms respectively).

In the case of the picture naming task, the highly proficient bilinguals outperformed the less proficient bilinguals. The highly proficient bilinguals were able to name the pictures more rapidly than the less proficient group when they were preceded by related and unrelated prime words. The size of the semantic interference effect for the highly proficient bilinguals was also significantly greater than that for the less proficient bilinguals. Analysis of the accuracy data showed that the highly proficient bilinguals were more accurate than the less proficient bilinguals across all conditions and tasks. These differences were significant across all conditions except for the related condition in the animacy decision task. The findings suggested that the highly proficient bilinguals were more sensitive to the manipulations of the stimuli types in both tasks relative to the less proficient bilinguals. It is relevant in the context of this study that both showed a similar pattern of performance in both tasks, which means the semantically related prime words affected their performance similarly, but that the size of the effect varied due to differences in their proficiency level.

c) Do script differences modulate the cross-language activation in both tasks?

As reported in section (a) the related primes induced a facilitation effect in the animacy decision task, contrasting with the picture naming task where they produced an interference effect. This suggests that script differences did not inhibit the conceptual processing of prime words in the animacy decision task, nor the conceptual and lexical processing in the picture naming task. In other words, the bilinguals processed the prime words although they were written in Arabic script. We replicated the findings of the semantic interference/facilitation in experiment one and two, confirming the non-selective access view in different script bilinguals. However, this required further investigation in experiments four and five, to determine whether this pattern of findings would persist across different experimental procurers. Regarding the manner of lexical selection, the findings of the primed picture naming task suggested that script differences did not modulate cross-language competition at the lexical level. This finding is in line with Boukadi et al. (2015), and our findings in experiment two (chapter five).

6.4 Conclusion

Attempting to find the locus of the semantic interference effect, the experiment described in this chapter compared the performance of highly and less proficient Arabic-English bilinguals in two tasks: the masked primed picture naming task (involving conceptual, lexical and sub-lexical processing) and the animacy decision task (involving conceptual processing). The lack of an interference effect in the conceptually based task suggested that the locus of the semantic interference effect is at the lexical level. The same pattern of findings was observed in both proficiency groups. With regard to the role of script in lexical access and manner of selection, experiment three provided evidence that script differences neither inhibit the activation of non-language across the different tasks, nor modulate cross-language competition for selection at the lexical level. We return to these findings in the general discussion (chapter nine).

Chapter 7 Experiment Four: Phoneme Monitoring Task I

The main objective of experiment four, the phoneme monitoring task, was to further investigate the manner of lexical access and the flow of activation during the process of word production in Arabic English bilinguals. Earlier chapters reviewed two opposing production models, namely the cascaded model (Dell 1986; Humphreys et al. 1988; Jescheniak and Schriefers 1998; Cutting and Ferreira 1999) which postulates that the activation of non-target lexical nodes spread to the phonological level, and the discrete model (Levelt 1989; Schriefers et al. 1990; Levelt et al. 1991; Levelt et al. 1999) which claims that lexical selection occurs first at the lexical level and then phonological information of the selected lexical node is activated at the phonological level. The conclusion that there is phonological activation of the non-target nodes at the phonological level came mainly from studies that adopted phoneme monitoring methodologically (Hermans et al. 1998; Colomé 2001) and naming cognate pictures tasks (Peterson and Savoy 1998; Janssen 1999). Here, in experiment one, we adopted the naming cognate task, but with modification to the priming paradigm. The results of experiment one suggested that the flow of activation of the non-target lexical nodes cascades to the phonological level, thereby challenging the claims of the discrete model. Therefore, it was decided to run a further experiment using another bilingual group of Arabic-English speakers, but adopting the phoneme monitoring task to establish whether the same pattern of findings is obtained with a different modality, thereby increasing validity. Recently, several studies (e.g., Moon and Jiang 2012) have speculated that cognate-based findings do not fully reflect the process of lexical access in bilinguals. They argued that cognate words only represent a small percentage of vocabularies, especially for different script bilinguals, so they cannot be considered representative of how lexical access occurs across all languages. They also assumed that cognates may enjoy special status in the bilingual lexicon, since they are lexically bilingual in nature; i.e., have a high level of form and semantic overlap. Thus, further investigation (experiment four) is required to validate our findings concerning the parallel activation of bilinguals' L1 and L2 languages at the phonological level.

The second objective of this experiment was to investigate further whether differences in L2 proficiency level would affect the manner of lexical access and the flow of activation; thus far, no effect has been identified in the previous masked priming experiments (cf chapters four, five and six). However, since the phoneme monitoring task is a different method from the cognate picture naming task and taps into the activation of phonological representations during lexical access, it was potentially informative to compare the performance of bilinguals with different proficiency levels (i.e., highly vs less) to determine whether the same findings could be obtained from different tasks. Moreover, studies (e.g., Colomé 2001) that adopted the phoneme monitoring task to investigate lexical access typically report on highly proficient bilinguals. Thus, data on less proficient bilinguals is lacking. We argue that more work is needed to determine whether differences in proficiency level modulate the flow of activation in bilinguals whose two languages have different/shared scripts.

As reviewed in chapter two (section 2.3.3), the phoneme monitoring task described has frequently been used to study the activation of those phonological representations involved in word production among bilinguals. The findings reported support the cascaded view (e.g., same script bilinguals: Colomé 2001; different script bilinguals: Moon and Jiang 2012). To the best of my knowledge, Moon and Jiang's (2012) study is the only one to date to have tested different script bilinguals by applying the phoneme monitoring task. Indeed, there are few published empirical studies have used this modified version of the phoneme monitoring task to investigate phonological activation in bilinguals. Therefore, experiment four not only investigated processing by a group of bilinguals that are rarely explored in bilingual lexical access studies, but it also contributes to the body of literature discussing the rarely employed phoneme monitoring task.

Adopting Moon and Jiang's (2012) method, we tested Arabic-English speakers in a phoneme monitoring task. We selected 14 phonemes that are similar in Arabic and English, in terms of place and manner of articulation (Alotaibi and Meftah 2013), and have a single grapheme representation in both languages. For example, the Arabic voiced bilabial stop /b/ is like the English /b/ as in *'**B**all.'* In both languages, the sound is formed by completely closing the lips, stopping the air flow in the oral cavity then releasing it. In term of its position in a word, it can occur *initially* (***B**ab* [English door]), medially (*Kabid* [English liver]), *or finally* (*Arnab* [English rabbit]).

Then we selected a set of pictures, whose labels in Arabic and English may or may not include the target phoneme. The task was monolingual; i.e., the participants were asked to respond whether the target phoneme was part of the target picture name in English as quickly as possible. The participants' responses were recorded to establish accuracy and response times. We constructed three conditions:

(i) The positive condition included pictures whose English names contained the target phoneme, and participants were expected to provide a positive response to the items in this condition;

(ii) The negative critical condition included pictures whose English names did not contain the target phoneme, but whose Arabic names did. The participants were expected to provide a negative response to the items in this condition;

(iii) The negative control in which the target phoneme was not part of the English or Arabic names of the target pictures, and the participants were expected to provide a negative response.

For example, a picture of a foot was preceded by the phoneme /f/ in the positive condition, and by the phoneme /r/ in the negative critical condition (which is part of the Arabic name for the picture i.e., *rejel* 'رجل') and by the 'random' phoneme /n/ in the negative control condition.

If the target and non-target lexical nodes are active at the phonological level, then longer naming latencies are anticipated in the negative critical condition relative to the negative control condition. In contrast, if only the target lexical node is active at the phonological level, then there should be no significant difference with the naming latencies between the negative critical and negative control condition. Thus, for the current research questions, we were particularly interested in comparing reaction times in the negative critical condition with the negative control condition. In addition, accuracy data was of equal importance here as we expected to observe more errors in the negative critical condition than for the negative control condition. If the two languages are active at the phonological level, then it would not only be hard to reject the items in the negative critical condition, but participants would also be more vulnerable to responding positively.

In short, this experiment allowed us to investigate whether the target and non-target lexical nodes across the two languages of the bilingual individual are co-activated, and whether this activation spreads to the phonological level. It also investigated the effect of proficiency level on the manner of lexical access, and the flow of activation across bilinguals whose languages have different scripts. Experiment four will enable us to examine cross-language phonological activation with zero phonological overlap between target and non-target names.

7.1 Methods

7.1.1 Participants

Ninety-eight Arabic-English adult bilinguals were recruited in this experiment, but of these, 12 were excluded due to technical errors. The majority of their responses (more than 60%) were not captured due to poor internet connection issues. Thus, in total, eighty-six participants took part fully in this experiment. All the participants had given consent to their data being used anonymously for research purposes prior to the study (see chapter three [section 3.2] for further details).

As with the other experiments, the participants were also divided into two groups (highly and less) according to their proficiency level (based on IELTS' scores), as detailed in chapter three (section 3.1). An independent-samples t-test was conducted to compare the IELTS scores between the two groups. This resulted in the identification of a statistically significant difference between the highly proficient group (M = 7.6, SD = .433) and the less proficient group (M = 4.16, SD = .810), t (70.7) = 25.58, p < .001. The results suggested a significant difference between the two groups in terms of their proficiency levels, as per their IELTS test scores. The results for these proficiency measures are discussed in the following section.

7.1.1.1 The results of the language history questionnaire

As summarised in Table 37, the similarities between the two groups were mainly related to their age of acquisition of L2 (which was around nine-years old). In addition, all participants received instruction in Arabic while being educated at elementary, intermediate and high schools. However, differences were recorded in terms of how they acquired English as a second language, and the language of instruction at university. Highly proficient participants primarily acquired English as a second language through their interactions with other people (mostly native English speakers as they pursued their postgraduate studies in native English-speaking countries), while the less proficient participants acquired English as a second language mainly through classroom teaching. With regard to the language of education at University (bachelor's degree level), the highly proficient bilinguals received instruction in English mainly as it was their major, whereas the less proficient bilinguals recruited from the institute of English Language received bilingual instruction, as some of the courses they were enrolled on were mainly Arabic based.

Moreover, the participants were asked to estimate their daily average use of L1 and L2. As shown in Table 37, the use of L1 and L2 by the less proficient group was 72.28% and 27.72%, respectively; whereas, the average estimation for the highly proficient group of L1 and L2 usage per day was 31.88 % and 68.13%, respectively. Table (37) shows the descriptive statistics for each question across the two groups.

The Highly Proficient Group (n=40)	**Minimum**	**Maximum**	**Mean**	**Std.Deviation**
Age (years)	30	40	35.98	2.29
IELTS	7	8	7.6	0.433
Mean daily L1 usage (%) (5 pt scale)	<25	50	31.88	12.64
Mean daily L2 usage (%) (5 pt scale)	50	100	68.13	12.64
Age of acquisition (years)	3	13	9.5	3.06

The Less Proficient Group (n=46)	**Minimum**	**Maximum**	**Mean**	**Std.Deviation**
Age (years)	19	22	20.15	1.13
IELTS	3	4	4.1	0.81
Mean daily L1 usage (%) (5 pt scale)	50	75	72.28	7.86
Mean daily L2 usage (%) (5 pt scale)	25	50	27.72	7.867
Age of acquisition (years)	3	13	9.5	2.8

Table. 37 Descriptive statistics for participants in both groups.

7.1.1.2 The results for the self-rating

The average scores for both groups in each skill are shown in Table 38. An independent-samples t-test was conducted to compare the mean scores for the self-assessed rating in each skill between the two groups, leading to the identification of statistically significant differences between the two groups in terms of all skill areas: (i) reading $t(80) = 21.362$, $p < .001$, (ii) writing $t(64) = 25.085$, $p < .001$, (iii) speaking $t(84) = 15.107$, $p < .001$, and listening $t(79) = 23.612$, $p < .001$. The results suggested that, based on their self-perception of their own proficiency level, the highly proficient bilinguals considered themselves as native-like, whereas the less proficient bilinguals considered themselves as beginners.

Group Statistics	Proficiency	N	Mean	Std. Deviation	Std. Error Mean
Self-Rate Reading (5 pt scale)	Highly	40	4.55	0.504	0.08
	Less	46	2.3	0.465	0.069
Self-Rate Writing (5 pt scale)	Highly	40	4.38	0.49	0.078
	Less	46	2.11	0.315	0.046
Self-Rate Speaking (5 pt scale)	Highly	40	4.5	0.555	0.088
	Less	46	2.5	0.658	0.097
Self-Rate Listening (5 pt scale)	Highly	40	4.58	0.501	0.079
	Less	46	2.13	0.453	0.067

Table. 38 Descriptive statistics for the self-assessed rating in participants' skills in L2 across the two groups.

7.1.1.3 The results of the lexical decision test

An independent-group t-test was performed on latencies and accuracy for words and nonwords, with proficiency (highly proficient bilinguals vs. less proficient bilinguals) identified as an independent variable. The highly proficient bilinguals were faster and more accurate at identifying nonwords than the less proficient bilinguals [$t(84) = 8.07$, $p = .000$ for latencies; $t(84) = -7.94$, $p = .000$ for accuracy]. Also the highly proficient bilinguals were significantly faster and more accurate with actual words than the less proficient bilinguals [$t(84) = 5.92$, $p = .000$ for latencies; $t(84) = -8.65$, $p =.000$ for accuracy]. These results suggest a significant difference between the two groups. The results of the lexical decision task for the highly and less proficient bilinguals are reported in Table 39.

	Non-word		**Word**	
Proficiency Groups	Reaction Times	Accuracy in percentage	Reaction	Accuracy in percentage
Highly proficient bilinguals (n-40)	1002 (196)	84.92 (11.3)	844 (156)	90.5 (3.6)
Less proficient bilinguals (n=46)	1529 (369)	67.93 (8.2)	1139 (278)	78.15 (8.3)

Table. 39 The results of the lexical decision task for the highly and less proficient bilinguals in experiment four. Standard deviations are in parentheses.

To conclude, with regard to the results of the language history questionnaire, the self-assessed rating and the lexical decision task suggested that the two groups differed significantly in terms of their proficiency level, which supported our system of classification based on their IELTS test scores.

7.1.2 Materials

Several steps were taken to construct the test materials for this study. Initially, 14 phonemes that are similar in English and Arabic were selected (see appendix J for the phonemes used for the experiment). Then forty-five pictures (see appendix H) were selected, and formed the stimuli list, and an additional set of twenty pictures were used as fillers (not included in the statistical analysis). The pictures were black and white line-drawings (see section 3.5.1) of animals, fruits, vegetables objects, furniture, etc., whose names in Arabic and English contained these phonemes. All pictures were non-cognates. We divided the stimuli list into three sets to create three experimental conditions:

(i) The positive condition consisted of 15 pictures, with English names containing the target phoneme (this constituted fillers in Moon and Jiang's (2012) study);

(ii) The control condition consisted of 15 pictures, whose English and Arabic names did not contain the target phoneme;

(iii) The critical condition consisted of 15 pictures, whose Arabic names contained the target phonemes.

To counterbalance the stimuli across conditions, we created three files to ensure that each picture appeared in all three conditions. For example, a picture of *an apple* appeared in the positive condition in file number one, in the negative condition in file two, and in the critical condition in file three. Thus, the participants were divided into three groups, with each group designated an experimental file. Furthermore, the frequency of the appearance of each phoneme was controlled, so that each phoneme appeared no more than 5 times among the 65 trials (stimuli and fillers). Although the participants were expected to monitor every syllable in the picture names, we ensured that the target phoneme appeared at the onset position in the critical and positive conditions to avoid the effect of different syllabic positioning. In each experimental file, we ensured that no more than two consecutive trials applied the same condition, and no more than four identical answers in a row.

All pictures were matched across the experimental conditions (see section 3.5.1). In brief, the picture names were controlled at $ps > .05$. Table 40 shows the average score for each variable along with the ANOVA results.

Variables	Stimuli set 1	Stimuli set 2	Stimuli set 3	*ANOVA Results*	***p*** **value (One-way ANOVA)**
Name Agreement (%)	.94	.93	.95	$F(2,42) = .795$	$p > 0.05$
Visual Complexity (KB)	17506	17435	17575	$F(2,42) = .001$	$p > 0.05$
Syllable Length	1.4	1.7	1.6	$F(2,42) = .350$	$p > 0.05$
Character Length	5.3	5.1	*5.5*	$F(2,42) = .365$	$p > 0.05$
Frequency per million words	3.4	*3.6*	*3.4*	$F(2,42) = .082$	$p > 0.05$
Age of Acquisition (1- 3 points scale)	2.13	*1.86*	*1.7*	$F(2,42) = .648$	$p > 0.05$
Familiarity (100-700)	546.7	554.43	541.9	$F(2,42) = .250$	$p > 0.05$
Imageability (100-700)	603.6	596.42	607.6	$F(2,42) = .465$	$p > 0.05$

Table. 40 Characteristics of pictures and picture names in English used in experiment four.

7.1.3 Design

The experiment was of a mixed design (i.e., a 2 x 3 factorial design) with proficiency level (highly proficient bilinguals and less proficient bilinguals) representing the between-subjects factors, and the picture-phoneme relationship representing the within-subjects factor (positive, critical and control).

7.1.4 Procedures

Each participant attended a video conference to meet with the researcher, and received oral instructions in English (see section 3.1.1.1.3) on how to effectively complete the task. The participants were trained individually on the phonetic symbols being used in the experiment. With the use of flash cards, explanations were first given regarding the sound of each of the 14 phonetic symbols (as indicated by English letters between a pair of slashes e.g., /t/), with a word from the target and non-target language as examples, but not the names of the pictures for use later in the experiment. We ensured that all the participants demonstrated perfect accuracy when deciding whether a word contained a target phoneme, as represented by a symbol, before moving on to the practice session, which was then followed by the test session.

The participants were randomly assigned to the three experimental files. They were asked to decide whether a target phoneme was part of the picture name in English or not. For each trial, inspired by Moon and Jiang (2012), the participants were first presented with a fixation sign (+) for 500 ms at the centre of the screen, which was then replaced with an English target phoneme that remained at the centre of the screen for 1000 ms. Then the phoneme was replaced with a picture of an object, food, etc. Pictures remained on the screen until the participants responded. If the phoneme was part of the English name of the picture, they were required to click (mouse click or screen touch click) the "yes" button on screen; if the target phoneme was not part of the English name of the picture, then they were required to click the "no" button on screen. After they responded, a fixation sign appeared for 500 ms again, and another target phoneme was presented. All the pictures were 430 by 430 pixels in size, and presented against a white background. The participants' responses were collected via a mouse click or touch screen clicks. Feedback was not provided for each response regarding response correctness. The task was created using PychoPy 3 described in chapter three (section 3.4.2).

7.2 Analysis of results

The data were collected from 86 participants. After trimming data (detailed in chapter three section 3.6.2), we then calculated reaction times and accuracy for correct responses again. Errors and outlier trials were excluded from the following analysis: (i) in the critical condition, error trials formed 13.88%, and outliers formed 0.2%, (ii) in the positive condition, error trials formed 5.80% and outliers 0.5%, and (iii) in the control condition, error trials formed 6.51% and outliers were 1.88 %. The mean RTs and percentage of accuracy are shown in Table 41.

Phoneme Type	Highly Proficient Bilinguals		Less Proficient Bilinguals	
	Reaction Times in ms	Accuracy in percentage	Reaction Times in ms	Accuracy in percentage
Positive condition	1106 (14)	96.65 (4.7)	1661 (34)	91.94 (8.7)
Critical condition	1815 (18)	93.23 (8.09)	2412 (37)	78.74 (9.2)
Control condition	1350 (20)	96.83 (6.40)	1955 (38)	90.26 (9.10)

Table. 41 Mean reaction times and percentage of accuracy for the less and highly proficient bilinguals. Standard deviations are in parentheses.

7.2.1 Reaction time analysis (RT)

A mixed ANOVA analysis was performed on mean response latencies per subject with, (i) proficiency level (highly proficient vs less proficient) as a between-subjects factor, and (ii) phoneme condition (positive, critical, and control) as a within-subjects factor. In addition, a repeated measure ANOVA was performed on the mean response latencies per item in each condition.

The results indicated a significant main effect of phoneme type in the analysis by subjects $F1(1, 83) = 569.24$, $MSE = 9.01$, $p < 0.001$, and in the analysis by items $F2(1, 88) = 245.57$, $MSE = 14.57$, $p < 0.001$. Also, as shown in Figure 21, there were significant differences in performance between the two groups. The main effect of proficiency type was significant in the analysis by subjects $F1(1, 83) =245.87$, $MSE = 21.81$, $p < 0.001$, and in the analysis by items $F2(1, 88) = 628.47$, MSE = 21.488, $p < 0.001$.

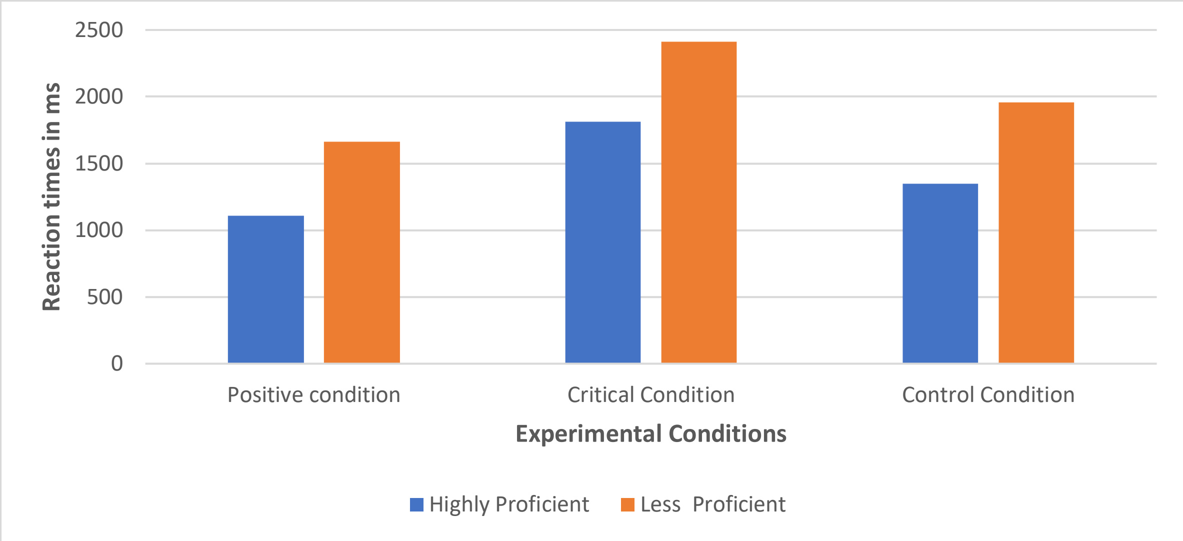


Figure. 21 Mean reaction times for the two proficiency groups across the three experimental conditions.

Moreover, the interaction between phoneme type and proficiency level was insignificant in the analysis by subjects $F1(1, 83) = 0.05$, $MSE = 0.001$, $p = 0.827$, and in the analysis by items $F2(1, 88) = 0.02$, $MSE = 0.001$, $p = 0.885$.

We conducted a paired t-test to examine the significance of the differences noted for phoneme type. As shown in Table 42, the average value for reaction times in the critical condition was significantly greater when compared to the control condition for the highly and less proficient groups ($p < 0.001$). However, the average value of reaction times for the positive condition was significantly smaller when compared to the control and critical conditions ($p < 0.001$) for the highly and less proficient groups. Finally, the average value of the reaction times for the critical condition was significantly larger in comparison with positive condition for both highly and less proficient groups. The calculated Cohen's d was > 0.8 for all the comparisons, indicating a "large" effect size for all conditions across participants.

Proficiency Level	Group 1	Group 2	Average Reaction Times (SE)		Difference between group 1 & group 2 (SE)	T-test (df)	p-value (Bonferroni correction)	Effect size (Cohen's d)	Effect size tabulation
			Group 1	Group 2					
High proficient bilinguals	critical	control	1815 (18)	1350 (20)	465 (24)	19.03 (39)	< 0.001	3.768	large
	positive	control	1106 (14)	1350 (20)	-244 (20)	-12.00 (39)	< 0.001	2.190	large
	critical	positive	1815 (18)	1106 (14)	709 (22)	31.61 (39)	< 0.001	6.663	large
Less proficient bilinguals	critical	control	2412 (37)	1955 (38)	456 (29)	15.62 (44)	< 0.001	1.780	large
	positive	control	1661 (34)	1955 (38)	-294 (24)	-11.90 (44)	< 0.001	1.191	large
	critical	positive	2412 (37)	1661 (34)	751 (31)	24.13 (44)	< 0.001	3.074	large

Table. 42 Paired t-test results to test the effect of phoneme type on RT across different Proficiency level.

To investigate the notable significant differences for proficiency level in all three conditions, we conducted a follow-up independent samples t-test. As summarised in Table 43, the average value of reaction times for the highly proficient participants was significantly smaller in all three conditions when compared with the less proficient participants.

Phoneme Type	Average Reaction Times (SE)		Difference between highly & less proficient bilinguals (SE)	T-test (df)	p-value	Effect size (Cohen's d)	Effect size tabulation
	Highly proficient Bilinguals	Less proficient Bilinguals					
Critical condition	1815 (18)	2412 (37)	-596 (42)	-14.12 (63.86)	< 0.001	1.47	large
Positive condition	1106 (14)	1661 (34)	-555 (38)	-14.62 (59.30)	< 0.001	1.52	large
Control condition	1350 (20)	1955 (38)	-605 (43)	-13.87 (66.00)	< 0.001	1.45	large

Table. 43 Independent samples t-test results to test differences in RT between high and less proficient groups across different experimental conditions.

7.2.2 Accuracy analysis (ACC)

A mixed ANOVA was conducted with proficiency level (highly and less) as a between-subjects factor and phoneme type (positive, critical and control) as a within-subjects factor. Results of mixed ANOVA showed a significant main effect of phoneme type in the analysis by subjects $F1(1, 83) = 50.50$, $MSE = 2420.74$, $p < 0.001$, and in the analysis by items $F2(1, 88) = 19.76$, $MSE = 2050.58$, $p < 0.001$. Also, the main effect of proficiency level was significant in the analysis by subjects $F1(1, 83) = 49.75$, $MSE = 4689.51$, $p < 0.001$, and in the analysis by items $F2(1, 88) = 66.73$, $MSE = 5333.09$, $p < 0.001$. Moreover, as shown in Figure 22, the interaction between phoneme type and proficiency level was significant in the analysis by subjects $F1(1, 83) = 13.80$, MSE = 661.83, $p = 0.002$, and in the analysis by items $F2(1, 88) = 6.23$, $MSE = 646.72$, $p = 0.014$. The corresponding eta-squared has a value of 0.143 (= 0.14), indicating a large effect size for these differences.

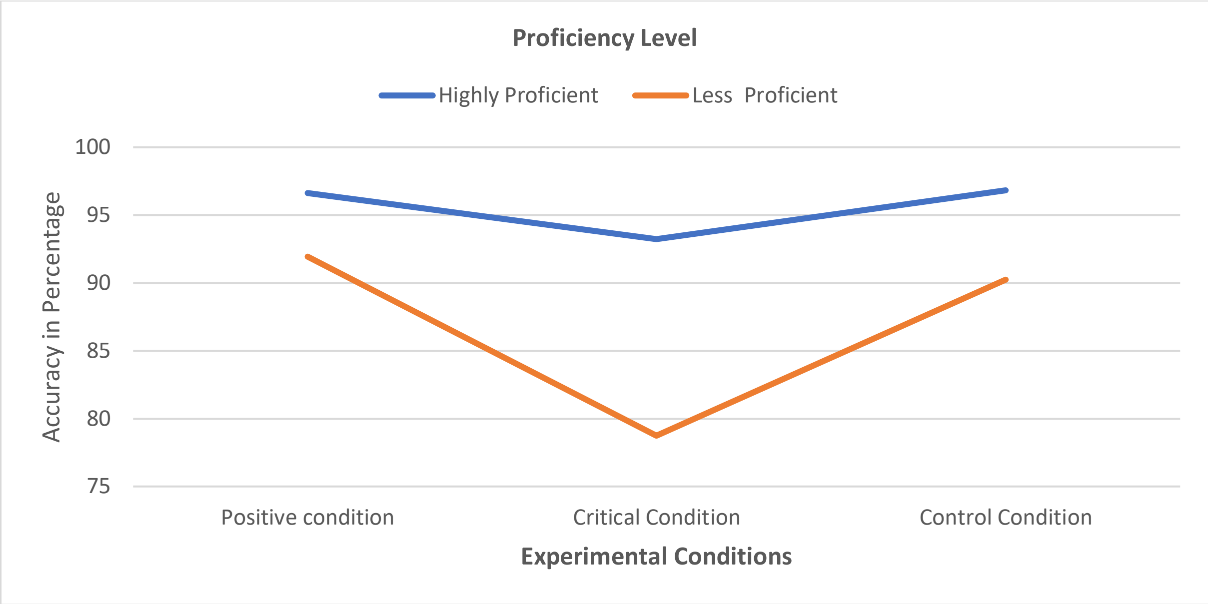


Figure. 22 Accuracy data in percentage for the less and highly proficient bilinguals across all experimental conditions.

To investigate the significant differences noted, a paired t-test was conducted. As shown in Table 44, the accuracy rate in the critical condition was significantly smaller when compared with the control condition for both the highly and less proficient groups. On the other hand, the accuracy rate in the positive condition was higher in comparison with the control condition for the less proficient group, and smaller for the highly proficient group; although these differences were insignificant.

Finally, the accuracy rate for the critical condition was significantly smaller compared with the positive condition for both the highly and less proficient groups. The differences in accuracy rate between the critical vs control condition, and between the critical vs positive condition were essentially larger (by magnitude) for the less proficient group, when compared to the highly proficient group.

Proficiency Level	Group 1	Group 2	Average Accuracy (SD)		Difference between group 1 & group 2 (SE)	T-test (df)	p-value	Effect size (Cohen's d)	Effect size tabulation
			Group 1	Group 2					
Highly proficient bilinguals	critical	control	93.23 (8.0)	96.83 (6.40)	-3.61 (1.39)	-2.59 (39)	0.027	0.498	medium
	positive	control	96.65 (4.7)	96.83 (6.40)	-0.18 (1.16)	-0.15 (39)	> 0.999	0.032	small
	critical	positive	93.23 (8.0)	96.65 (4.7)	-3.43 (1.42)	-2.41 (39)	0.041	0.532	medium
Less proficient bilinguals	critical	control	78.74 (9.2)	90.26 (9.19)	-11.51 (1.58)	-7.29 (44)	< 0.001	1.257	medium
	positive	control	91.94 (8.7)	90.26 (9.19)	1.68 (1.59)	1.06 (44)	0.594	0.188	small
	critical	positive	78.74 (9.2)	91.94 (8.7)	-13.19 (1.71)	7.74 (44)	< 0.001	1.467	medium

Table. 44 Paired t-test results to test the effect of phoneme type on ACC (in percentage) across different Proficiency subgroups.

Since there was a significant main effect of proficiency level, we conducted an independent sample t-test. As shown in Table 45, the highly proficient group was significantly more accurate than the less proficient group across all the experimental conditions. A significant interaction between phoneme type and proficiency level in accuracy rate was also identified, which indicated that differences in accuracy between high and less proficient bilinguals arose due to phoneme type. Considering the data in Table 45, it was apparent that the difference between the less proficient bilinguals and the highly proficient bilinguals was large in magnitude in the critical conditions compared with the other experimental conditions. The calculated Cohen’s d was = 0.8, indicating a large effect size.

Phoneme Type	Average Accuracy (SD)		Difference between highly & less proficient bilinguals (SE)	t-test (df)	p-value of *t*-test	Effect size (Cohen's d)	Effect size tabulation
	Highly proficient bilinguals	Less proficient bilinguals					
Critical condition	93.23 (8.0)	78.74 (9.2)	14.48 (1.89)	7.66 (83)	< 0.001	0.83	large
Positive condition	96.65 (4.7)	91.94 (8.7)	4.72 (1.51)	3.12 (69.56)	0.003	0.33	between small and middle
Control condition	96.83 (6.40)	90.26 (9.19)	6.58 (1.69)	3.88 (78.99)	< 0.001	0.41	middle

Table. 45 Independent samples t-test results to test the differences in ACC (in percentage) between the high and less proficient groups across all the experimental conditions.

7.3 Discussion

a. Is the flow of activation in different script bilinguals cascaded or discrete?

The results of this experiment showed that there was a significant main effect of phoneme type. This indicated that the participants' performance in the three experimental conditions (i.e., control, critical and positive condition) differed. When comparing the naming latencies between all three conditions, we found that the phonemes in the critical conditions yielded longer naming latencies (a difference of 465 ms) relative to the control condition. By contrast, phonemes in the positive conditions produced faster reaction times (a difference of 265 ms) relative to the control condition. The results demonstrated that although the task was based on monitoring phonemes for the English name of the picture only, the phonological representations of the non-target lexical nodes (in Arabic) are active at the phonological level, as the participants were slower at rejecting them. The analysis of accuracy data showed more errors in the critical condition relative to the control condition (an estimated difference of 15.12%). In addition, a high rate of accuracy was observed in the positive condition relative to the control condition (a difference of 1.86 %). As hypothesized, if the two languages are active at the phonological level, it would be difficult to reject the items in the critical condition, and the participants are more vulnerable to responding positively. The findings here support the cascaded view (Dell 1986; Cutting and Ferreira 1999) and are also in accordance with Moon and Jiang's study (2012), in which significant differences were observed between the critical and control condition and between the positive and the control condition when testing Korean-English speakers. The findings suggest phonological co-activation is automatic for different script bilinguals, even when there is no phonological/orthographical overlap between the target and non-target names. The fact that the two languages (Arabic-English) have different scripts did not inhibit activation of the non-target name. Experiment four extended the finding when naming cognates (experiment one, chapter four), such that even when the two alternatives were not phonologically similar, activation of the target and non-target name was inevitable. The findings reported from experiment four indicate that script differences play no particular role in bilingual lexical access. In the next chapter, we investigate this issue further.

However, the results of this experiment were not in accordance with Hermans et al. (2011), as although all the filler items in this experiment were not cognates, cross language phonological activation was observed, unlike in their study. A possible reason for this contradictory result could be due to differences in the characteristics of the selected materials. The appearance of phonemes in experiment four was controlled and the characteristics of the stimuli lists were matched; whereas, in Hermans et al.'s (2011) study (third experiment), some phonemes appeared more frequently than others did. In fact, the dominant presence of certain phonemes was apparent in the filler and experimental item lists, which might have skewed their results. To illustrate, two phonemes /t/ and /b/ had a dominant presence in specific conditions. The phoneme /t/ appeared 8 times in the positive condition, but had zero occurrence in the critical and unrelated conditions, while the phoneme /b/ was very dominant in the critical and control conditions, compared to the positive condition. Therefore, having a dominant phoneme in a certain condition could influence the performance of the participants i.e., seeing the /t/ phoneme promoted a positive response to the target picture, whereas seeing the /b/ phoneme promoted a negative response. This correlation could facilitate a response and thus reduce the difference in naming latencies between the two conditions.

b- What is the effect of proficiency level on the cross-language phonological activation?

With regard to proficiency level, the results of the reaction times analysis revealed a significant main effect of proficiency level indicating significant differences between the performance of the highly proficient and less proficient participants. The statistical comparative analysis revealed that highly proficient bilinguals were faster than less proficient bilinguals in all three experimental conditions. Specifically, they were 600 ms faster in the critical condition, 560 ms faster in the positive condition, and 610 ms faster in the control condition. Moreover, the results of the accuracy analysis demonstrated that the highly proficient bilinguals were more accurate than the less proficient bilinguals in the critical, positive and control conditions; i.e., estimated differences of 14.48%, 4.72%, 6.58% respectively. The findings pertaining to enhanced performance for highly proficient bilinguals correspond to experiments one, two and three (chapters four, five and six, respectively), and again support the Revised Hierarchical Model predictions regarding impediments to the performance of less proficient bilinguals.

Moreover, these findings suggest the non-target language is non-selectively activated during production, regardless of the bilinguals' proficiency level. The fact that both groups took longer to reject the L1 phoneme in the critical condition, although they were asked specifically to decide whether the English name of the picture had the target phoneme, demonstrates that L1 was active and formed some sort of interference. This finding is in line with that of Moon and Jiang (2012) who tested highly proficient Korean-English bilinguals in a phoneme monitoring task, and reported that highly proficient bilinguals experience non-selective access, and that activation cascaded to the phonological level.

In addition, an interesting finding was that a significant interaction between phoneme type and proficiency level in accuracy rate was evident, which implied that differences in accuracy between highly and less proficient bilinguals were affected by phoneme type and we will evaluate further this finding in experiment five (chapter eight). In the critical condition, the magnitude of the differences between the less proficient bilinguals and the highly proficient bilinguals was notably large compared to other conditions. The less proficient participants were more vulnerable to replying positively when the phoneme was part of the picture name in L1, than the highly proficient bilinguals were. This finding can be explained in line with the Revised Hierarchical Model (RHM) proposed by Kroll and Stewart in 1994 (described in section 2.4.2). This model assumes that L2 learners retrieve/access the concept of L2 lexical nodes through L1 lexical nodes (i.e., translation). As proficiency level increases, links between L2 and concept become stronger and no reliance on L1 lexical links is needed. Considering this, we postulated that the less proficient bilinguals in this experiment were more error-prone than the highly proficient bilinguals, because for the less proficient bilinguals, the L2 name of the picture is strongly connected to the L1 name of the picture, and thus more interference from L1 nodes is predicted. Moreover in (2016), Jacobs et al. tested the effect of cross language cognates (phonologically related words) on the production of L2 words, and found that the less proficient bilinguals experienced more difficulty inhibiting cross-linguistic activation patterns during speech production than the highly proficient bilinguals. As a consequence, the influence of the non-target lexical nodes spilled over into the articulatory realization of the phonologically related words. Therefore, at the word level, the less proficient bilinguals were unable to control the interference of the activated phonologically-related L1 nodes, and seemingly this lack of control exists even when the interference is caused by a single phonologically- related phoneme, as the accuracy data indicated in this experiment.

Prior to making any generalisation, we compared the findings of this experiment with other studies in the literature that tested the effect of proficiency level on cross linguistic activation in bilingual word production. In (2006), Schwartz and Kroll compared the performance of highly proficient Spanish-English speakers with the intermediate proficiency Spanish-English bilinguals in an L2 reading task. They examined the nature of cross-language lexical competition in the sentence context for two groups of bilinguals. Specifically, they investigated whether L1 (non-target language) would interfere when individuals were engaging in L2 (target language) processing.

First, the reaction times data showed minimal interference in the critical condition compared to the control condition, whereas accuracy data showed significant differences between the critical and the control conditions. Of relevance, no significant differences were noted between the two groups in their reaction times; however, the accuracy date revealed significant differences between the two groups. They postulated that the similarities in reaction times between the two groups were due to: (i) the nature of the task being executed over a relatively short period of time, and thus being less likely to reflect general differences in proficiency; and (ii) the large number of errors that were found, (not included in the calculation of RT) reduced the average reaction time for the intermediate group; and (iii) the intermediate proficiency bilinguals were fairly competent in L2, and capable of understanding and performing the task.

Second, a higher error rate was found for the less proficient group relative to the highly proficient group. They attributed the differences in accuracy to the fact that the highly proficient bilinguals had been exposed to a minimal activation of the non-target lexical nodes during processing in the L2, unlike the less proficient bilinguals. They concluded that as proficiency increases, L1 is less likely to be active or interfere in an L2 processing tasks.

When comparing the results of the two studies, the chief difference between the current study and Schwartz and Kroll's (2006) was that the performance of the two groups of bilinguals did not differ significantly in their study, with regard to the reaction times data, whereas in this experiment, the two groups of Arabic English bilinguals differed significantly. A possible reason for this discrepancy is that the size of the difference in proficiency level between bilinguals in Schwartz and Kroll's (2006) study was small compared to that in the present experiment. They tested intermediate and highly proficient bilinguals, whereas in this experiment we tested less proficient (i.e., L2 learners) and highly proficient bilinguals (L2 advanced speakers). Adjacency to their proficiency level, such as variations in performance did not stand out. Hence, differences between them did not reach a statistically significant level. Moreover, the sample size in this experiment exceeded that of Schwartz and Kroll (2006), and thus more data were generated for the purpose of inclusion in the analysis. A third possible reason for the discrepancy in findings can be attributed to the task differences and the processing stages involved. Schwartz and Kroll's (2006) study utilised a form-related competitor in a sentence processing task, requiring rich semantic processing, unlike the phoneme monitoring task which involves superficial form processing only. Conversely, the findings from the accuracy data for both studies were similar, and supported the theory that L1 is active and affects the performance of both groups during L2 processing.

It is important to consider that in experiment four, shared phonemes across the two languages were used to investigate cross-language activation and to which level this activation penetrates. Thus, we believe there is a need to explore whether the same pattern of findings is obtained when a phoneme is not shared across the two languages. Arabic for example has phonemes that do not exist in English, so we wished to investigate whether presenting phonemes that are part of the L1 but not L2 picture names would induce longer naming latencies and lower rates of accuracy or not. The final experiment (five) in this thesis addressed this question, and compared the performance of the participants across two different conditions, shared phonemes and distinct phonemes.

7.4 Conclusion

Experiment four was conducted to investigate the manner of cross-language activation during single word production in bilinguals using the phoneme monitoring task. More specifically, it examined whether the target and non-target lexical nodes across the two languages (Arabic-English) are co-activated in bilinguals, and whether this activation spreads to the phonological level; i.e., whether the flow of activation cascades to the lexical level only (discrete model) or to the phonological level (cascaded model). It also investigated the effect of proficiency level on the cross-language activations of different script bilinguals. The results showed that cross-language activation extends to the phonological level in the case of different script bilinguals. This finding supports those of previous studies that have examined the co-activation of phonological presentation at the phonological level using phoneme monitoring tasks. However, as we mentioned earlier, the results cannot be generalized, as only phonemes shared across the two languages were used here. Thus, in the next chapter, we will expand this experiment to include language specific phonemes, and to determine whether we obtain the same pattern of findings or not. In terms of proficiency level, we conclude that this did not modulate cross-language activation, as the same pattern of findings emerged for both groups.

Chapter 8 Experiment Five: Phoneme monitoring Part II

In experiment four, we investigated lexical access and the flow of activation during single word production in bilinguals using the phoneme monitoring task. We presented bilinguals with phonemes that are similar in the Arabic and English language, followed by a picture of an object. The participants had to retrieve the name of the picture, and decide whether the phoneme displayed was part of the English (i.e., L2) name of the picture. We manipulated the type of the phoneme presented; i.e., a phoneme that was part of the L2 picture name, a phoneme that was part of the L1 picture name and a phoneme not present in either name. The results showed that it was difficult to reject the phoneme when it was part of the L1 picture name. We concluded that L1 was activated at the phonological level, and that proficiency level did not modulate this cross-language activation. However, this finding cannot be generalized because we only tested part of the phonological system, namely phonemes common to both languages, and we hypothesize that distinct phonemes will be processed differently. According to Simonet (2016), the majority of phonological forms in bilinguals' two languages are shared; in other words, the phonological segments are reutilized. However, it is posited (Flege 1987) that when phonological forms are not shared, then bilinguals will develop links between similar sounds across the two languages.

Therefore, for example, they establish a connection (backward link) between the phonetic category in the L2 language (the new language) and the closest similar phonetic category from the L1 language. To illustrate, the two English phonemes /p/ and /b/ are considered to be distinct phonemes for native English speakers, but for early Arabic-English learners, they sound similar. Thus, those learners often produce an Arabic phoneme /b/ (i.e., a phonetically similar phoneme) to represent the two English phonemes /b/ and /p/. In other words, they create backward links from L2 distinct phonemes to L1 similar phonemes. Consequently, it is logical to propose that, since Arabic distinct phonemes do not exist in English, there are no shared links between these phonemes and any similar phonemes in English. In summary, there is no opportunity to create forward links from Arabic distinct phonemes to similar English phonemes.

However, if such links do exist, they would be expected to be much weaker than backward links, i.e., from English to Arabic. The only possible scenario for Arabic specific sounds to occur in an English conversation involving an Arabic-English bilingual, for example, is when mentioning the Arabic names of persons, places, …etc. Thus, we can assume that the Arabic distinct phonemes rarely occur in English contexts. Therefore, we hypothesized that during (English) production, Arabic distinct phonemes would not be expected to be triggered by the activation of any English phonemes, unlike the shared phonemes across the two languages. Shared phonemes are highly active across the two languages, as they are likely to be produced more frequently, whereas language distinct phonemes are produced in one language only. Thus, in the L2 production task, shared phonemes across the two languages are more active for a bilingual speaker than L1 specific phonemes. In addition, distinct Arabic script may serve to inhibit the activation of language specific phonemes, as they typically appear in an Arabic context. Shared phonemes have zero orthographic overlap, yet share the same phonotactic features; whereas, in contrast distinct phonemes have zero orthographic and phonological overlap. Therefore, the possibilities of co-activation when presenting special phonemes are low.

This leads us to question how this may challenge the findings of phoneme monitoring task. We have reported in chapter seven (experiment four) that presenting a shared phoneme that is highly active in two languages triggered the activation of both languages. It is possible that the presence of a shared phoneme pushes the two languages towards a standby position. On the other hand, presenting an L1 specific phoneme, which has a slight likelihood of occurrence in an L2 task is unexpected, and thus the normal reaction is to totally ignore it or reject it, since the participants were instructed to retrieve the English name of the picture. It is easier to reject a phoneme that is L1 distinct in this task than a phoneme that is shared across two languages. Therefore, we expected a rejection decision to be made once the L1 specific phoneme is presented; i.e., prior to the presentation of the picture. This differs from the shared phoneme that cannot be rejected unless a picture is presented, and its name is retrieved. Thus, identifying an interference effect in that condition suggests the name of the picture is simultaneously activated in the two languages regardless of the phoneme type.

We replicated the same experimental conditions (with different groups of participants) as those used in our previous phoneme monitoring task with the addition of one more critical condition. That is, for experiment five we created four conditions: the control condition, the positive condition, the bilingual critical condition (shared phonemes) and the monolingual critical condition (language distinct phonemes). This new monolingual critical condition should assist when determining whether or not the finding of cross language phonological activation was limited to using phonologically similar phonemes from both languages. Unlike experiment four, we ensured that all L1 picture names had language distinct phonemes. This allowed us to use the same picture for the two critical conditions, which enabled us to investigate whether or not the L1 name of the picture was activated due to the type of the phoneme presented; i.e., a shared or L1 specific phoneme. For example, the picture of a *spider* ‘ عنكبوت ’ was preceded by the phoneme /n/ in the bilingual critical condition; whereas, in the monolingual critical condition, it was preceded by the phoneme /ʕ/ ع . There were two possible outcomes identified to describe the monolingual critical condition:

(i) Slower reaction times that are similar in size to the bilingual critical condition. This would indicate that target and non-target lexical nodes were activated at the phonological level, regardless of phoneme type. This would provide support to a fully cascaded activation of the non-target nodes.

(ii) Faster reaction times relative to the control condition (i.e., unrelated phoneme to L1 or L2 picture name), which suggest it would be easier to reject the L1 phoneme due to its distinct phonological features. Thus, this would indicate that the cross-language phonological activation seen in the phoneme monitoring task occurred due to the use of shared phonemes across the two languages. Thus, a fully cascaded activation of the non-target nodes at the phonological level cannot be assumed.

For experiment five, we have identified eight distinct Arabic phonemes. These are all consonants with no equivalents in English: pharyngeal phonemes (unvoiced fricative /ħ/ ح and unvoiced fricative /ʕ/ ع), two uvular fricative phonemes (unvoiced non-emphatic /χ/ خ and voiced non-emphatic /ɣ/ غ) , one intro-dental phoneme (voiced emphatic /ðˤ/ ظ) and alveo-dental (voiced emphatic /dʕ/ ض , unvoiced emphatic /tʕ/ط , unvoiced emphatic /sʕ/, ص) (AlMahmoud 2013). These Arabic consonants typically present considerable difficulties for English learners, due to their phonotactic differences from English (Rammuny 1976; Alosh 1987; Shehata 2015). For a complete list of these phonemes see appendix J.

8.1 Methods

8.1.1 Participants

A new group of participants were recruited for this experiment, who had not taken participated in any other task related to this study. 112 Arabic-English adult bilinguals participated, but 20 were excluded due to a technical error; i.e., more than 50% of trials were not captured due to internet connection issues, and six participants were excluded because they had aborted the task before allowing the software to save their results. Therefore, in total, eighty-six participants took part fully in this experiment. All the participants had given consent to their data being used anonymously for research purposes prior to the study (see chapter three for further details).

Since all the participants had taken the International English Language Testing system test (IELTS), we used the test scores as the primary measure of proficiency to validate our classifications of the two groups required in the study. An independent-samples t-test was conducted to compare IELTS scores between the two groups. This resulted in the identification of a statistically significant difference in the IELTS test scores between the highly proficient group (M = 7.9, SD = .53) and the less proficient group (M = 4.3, SD = .77), t (79) = 24.14, p < .001.

8.1.1.1 The results of the language history questionnaire

As summarised in Table 48, the similarities between the two groups were mainly related to their age of acquisition of L2 (which was around nine-years old), and the fact that they had received instruction in Arabic during their education at elementary, intermediate and high school.

However, differences emerged regarding how they acquired English as a second language, and their language of instruction at university. The highly proficient participants had primarily acquired English as a second language through interacting with other people (mostly native English speakers as they pursued their postgraduate studies in native English-speaking countries). However, the less proficient participants had acquired English as a second language mainly through classroom teaching.

With regard to the language of education at University (bachelor's degree level), the highly proficient bilinguals received instruction in English mainly, as it was their major; whereas, the less proficient bilinguals, recruited from the institute of English Language, received instruction in the bilingual mode, as some of the courses they were enrolled on were mainly Arabic based. Moreover, the participants were asked to estimate their daily average use of L1 and L2. As shown in Table 46, the use of L1 and L2 by the less proficient group was 72.83% and 27.17%, respectively, while the average estimation for the highly proficient group of L1 and L2 usage per day was 36.88 % and 63.13% respectively. Table 46 shows the descriptive statistics for each question for the two groups. Overall, the less and highly proficient groups were closely matched in terms of L2 age of acquisition (AoA).

The Highly Proficient Group (n= 40)	**Minimum**	**Maximum**	**Mean**	**Std.Deviation**
Age (years)	30	40	36.17	2.286
IELTS	7	9	7.9	0.533
Mean daily L1 usage (%) (5 pt scale)	< 25	50	36.88	13.853
Mean daily L2 usage (%) (5 pt scale)	50	100	63.13	13.853
Age of acquisition (years)	3	13	9.5	3.063
The Less Proficient Group (n= 46)	**Minimum**	**Maximum**	**Mean**	**Std.Deviation**
Age (years)	19	22	20.15	1.135
IELTS	3	4	4.3	0.775
Mean daily L1 usage (%) (5 pt scale)	50	100	72.83	13.77
Mean daily L2 usage (%) (5 pt scale)	< 25	50	27.17	13.77
Age of acquisition (years)	3	13	9.5	2.802

Table. 46 Descriptive statistics for participants in both groups.

8.1.1.2 The results for the self-rating

The average scores for both groups for each skill are shown in Table 47. We conducted an independent-samples t-test to comparing the mean scores for the self-assessed rating in each skill area between the two groups. The results showed statistically significant differences between the two groups for all skills: (i) Reading $t(84) = -19.20$, $p < .001$, (ii) writing $t(61.29) = -24.10$, $p < .001$, (iii) speaking $t(77) = -18.96$, $p < .001$, and listening $t\ (79) = -\ 23.61$, $p < .001$. These results suggest that, based on their own perception of their proficiency level, the highly proficient bilinguals considered themselves to be native-like; whereas, the less proficient bilinguals considered themselves as beginners.

Skills	Proficiency Level	N	Mean	Std. Deviation	Std. Error Mean
Self-Rated Reading (5 pt scale)	Less Proficient	46	46	2.15	0.47
	Highly Proficient	40	40	4.25	0.543
Self-Rated Writing (5 pt scale)	Less Proficient	46	46	2.07	0.327
	Highly proficient	40	40	4.47	0.554
Self-Rated Speaking (5 pt scale)	Less Proficient	46	46	2.54	0.657
	Highly proficient	40	40	4.78	0.423
Self-Rated Listening (5 pt scale)	Less Proficient	46	46	2.37	0.572
	Highly proficient	40	40	4.83	0.385

Table. 47 Descriptive statistics for the self-assessed rating in participants' skills in L2 across the two groups.

8.1.1.3 The results of the lexical decision test

An independent-group t-test was performed on latencies and accuracy for words and nonwords with proficiency (highly proficient bilinguals vs. less proficient bilinguals) as the independent variable. The performance of the highly proficient bilinguals and the less proficient bilinguals differed significantly terms of accuracy and reaction times for word and nonwords (Table 48).

In the non-word condition, the highly proficient bilinguals were faster and more accurate than the less proficient bilinguals [$t(84) = 6.61$, $p < .001$ for latencies; $t(84) = -10.84$, $p < .001$ for accuracy]. Additionally, in the word condition, the highly proficient bilinguals were significantly faster and more accurate than the less proficient bilinguals [$t(84) = 7.33$, $p < .001$ for latencies; $t(65.9) = -11.26$, $p < .001$ for accuracy]. These results suggested a significant difference between the two groups.

	Non-word		**Word**	
Proficiency Groups	Reaction Time	Accuracy in percentage	Reaction Time	Accuracy in percentage
Highly Proficient Bilinguals (n=40)	1099 (332)	87 (8.1)	783 (136)	93 (3.4)
Less Proficient Bilinguals (n=46)	1596 (359)	68 (8.2)	1135 (276)	79.06 (7.3)

Table. 48 The results for the lexical decision task for the highly and less proficient bilinguals in experiment five. Standard deviations are in parentheses.

In conclusion, the results of the language history questionnaire, the self-assessed rating and the lexical decision task suggested the two groups differed significantly in their proficiency level, which supported our classification that was based on their IELTS test scores.

8.1.2 Materials

To construct the test materials for this experiment, we first identified the Arabic distinct phonemes. As stated previously, we selected eight phonemes: /ħ/ ح, /ʕ/ع , /χ/ خ , /ɣ/ غ , /dʕ/ ض , /tʕ/ط , /sʕ/, ص, and /ðʕ/ ظ to create the monolingual critical condition, and we adopted the 14 similar phonemes for use in experiment four to create the bilingual critical condition (see appendix J for a description of these phonemes).

After this, we selected stimuli pictures which had names in Arabic containing at least one distinct phoneme (for a full list of these picture names, see appendix I). Eighty-six pictures were selected for this experiment: sixty pictures formed the stimuli list and an additional set of 26 pictures were used as fillers and for practice trials (20 and 6 respectively). The selected pictures represented objects including home furniture, vegetables, fruits, animals, cloths, tools, musical instruments...etc. (for details see section 3.5.1). All the stimuli pictures had at least one distinct phoneme in their Arabic name, and had non-cognate names. We matched pictures and picture name across conditions on several variables (for details see section 3.5.1). In brief, the picture names were controlled at *p*s > .05. Table 49 shows the average score for each variable along with the ANOVA results.

Variables	***Stimuli set 1***	***Stimuli set 2***	***Stimuli set 3***	***Stimuli set 4***	***ANOVA Results***	***p value (One-way ANOVA)***
Name Agreement (%)	0.86996	0.94537	0.89018	0.89488	*F(3,56) =.1.159*	*p > 0.05*
Visual Complexity (KB)	15707	13486	16263	17282	*F(3,56) = .559*	*p > 0.05*
Syllable Length	1.4	1.4	1.67	1.47	*F(3,56) = .601*	*p > 0.05*
Character Length	4.87	5.27	5.73	4.93	*F(3,56) = 1.077*	*p > 0.05*
Frequency per million words	2.83	3.23	2.96	2.78	*F(3,56)= .400*	*p > 0.05*
Age of Acquisition (1- 3 points scale)	2.07	2	2.33	2.47	*F(3,56) = .886*	*p > 0.05*
Familiarity (100-700)	565.5	548.79	504.73	536.93	*F(3,56) = 2.36*	*p > 0.05*
Imageability (100-700)	597	582.5	596.1	575	*F(3,56) = .737*	*p > 0.05*

Table. 49 Characteristics of pictures and picture names in English across the four experimental conditions.

To avoid the effect of different syllabic positioning, we tried to maintain the onset position for the distinct phonemes in all picture names. Unfortunately, we could not apply this to all stimuli lists, since the picture names had to be controlled on multiple levels. Thus, the distinct phonemes, in the monolingual critical condition, appeared in the first syllable and in the onset position for 84% of the target picture names (e.g., طبلة /tˤab.la/ [English drum]), while for the remainder they appeared at the second syllable (e.g., قطار /qi.tˤaː.r/ [English train]). The shared phonemes, in the bilingual critical condition, were always in the second or the third syllable. Since our main objective was to observe the effect of using a distinct phoneme in this task, we prioritized the position of the distinct phonemes over the shared phonemes when selecting the stimuli pictures. Furthermore, the participants were expected to monitor every syllable in the pictures' names and thus, we assumed that the impact of word position was insignificant.

We divided the stimuli list into four sets to create four experimental conditions:

(i) The control condition which consisted of 15 pictures whose English and Arabic names did not contain the target phoneme;

(ii) The positive condition which consisted of 15 pictures whose English names only contained the target phoneme;

(iii) The monolingual critical condition which consisted of 15 pictures whose Arabic names contained the language-specific target phonemes;

(iv) The bilingual critical condition which consisted of 15 pictures whose Arabic names contained the target phoneme.

Then by creating four experimental files, we counterbalanced the stimuli across the conditions to ensure that each picture appeared in all four conditions. For example, a picture of *grapes* " عنب " appeared in the positive condition preceded by the phoneme /g/ in file one, in the control condition preceded by the phoneme /l/ in file two, in the monolingual critical condition preceded by the phoneme /ʕ/ع in file three, and in the bilingual critical condition preceded by the phoneme /n/ in file four. The within file constrictions were:

(i) Shared phonemes which appeared no more than 7 times across the ninety trials (i.e., for stimuli and fillers);

(ii) The distinct phonemes which appeared no more than 6 times across the ninety trials;

(iii) No more than two consecutive trials applied the same condition;

(iv) No more than four identical answers (i.e., yes or no) in a row.

To avoid an order effect, we randomized the order of the trials across the four experimental files. The participants were also randomly divided into the four groups; each group had an equal number of highly proficient and less proficient participants and each was assigned to an experimental file.

8.1.3 Design

The experiment applied a mixed design (i.e., a 2 x 4 factorial design) with proficiency level (highly proficient participants and less proficient participants) representing the between-subjects factor and picture-phoneme relationship representing the within-subjects factor with four levels (positive, monolingual critical, bilingual critical and control).

8.1.4 Procedures

The participants were trained individually on the phonetic symbols that would be used in the experiment. With the use of flash cards, explanations were first given regarding the sound of each of the twenty two phonetic symbols (English graphemes put in a pair of slashes for shared phonemes e.g., /t/, and specific phonetic symbols to represent the Arabic distinct phonemes /ħ/), with a few words from the target language L2 given as examples for shared phonemes, but not the names of the pictures that would be used later in the experiment. Before moving to the practice session on Psychopy, we ensured that all the participants demonstrated perfect accuracy when recognizing the phonemes. The same procedures that were applied in experiment three were used here and have been described in chapter three. The participants had to decide whether or not the visually presented phoneme is part of the L2 picture name.

Ethical approval was obtained from the ENCAP Research Ethics Committee at Cardiff University, and prior to the experiment, written informed consent was obtained from all the participants.

8.2 Analysis of Results

The data were collected from eighty-six participants. After trimming the reaction time data (detailed in section 3.6.2), we calculated reaction times and accuracy for correct responses again. Errors and outlier trials were excluded from the following analysis:

(i) In the monolingual critical condition, the error trials formed 19.84% and the outliers formed 1.71%;

(ii) In the bilingual critical condition, the error trials formed 17.05% and the outliers 1.86%;

(iii) In the control condition, the error trials formed 7.29% and the outliers were 0.85%;

(iv) In the positive condition, the error trials formed 6.20% and the outliers 0.23%.

The mean reaction times and percentage of accuracy are shown in Table 50.

Phoneme Type	Highly proficient bilinguals		Less proficient bilinguals	
	Reaction times in ms	Accuracy in percentage	Reaction times in ms	Accuracy in percentage
Positive condition	1176 (48)	98.00 (3.4)	1664 (32)	90.11 (7.3)
Monolingual critical condition	1892 (59)	84.81 (14. 9)	2564 (37)	77.21 (9.7)
Bilingual critical condition	1882 (55)	87.21 (10.7)	2388 (31)	79.03 (10.7)
control condition	1463 (56)	96.83 (6.3)	2061 (36)	89.49 (9.4)

Table. 50 Mean reaction times and percentage of accuracy for the less and highly proficient bilinguals. Standard deviations are in parentheses.

8.2.1 Reaction times analysis (RT)

A mixed ANOVA analysis was performed on the mean response latencies per participant, with (i) proficiency level (highly proficient vs less proficient) as the between-subjects factor, and (ii) phoneme condition in (positive, monolingual critical, bilingual critical and control) as the within-subjects factor. In addition, a repeated measure ANOVA was performed on the mean response latencies per item for each condition.

The results indicated a significant main effect of phoneme type in the analysis by subjects $F1$ (1, 84) = 242.012, MSE = 5.36, p < 0.001, and in the analysis by items $F2$ (1, 118) = 95.38, MSE = 7.45, p < 0.001. In addition, the main effect of proficiency type was significant in the analysis by subjects s$F1$(1, 84) = 106.17, MSE = 27.42, p < 0.001, and in the analysis by items $F2$ (1, 118) = 528.73, MSE = 36.168, p < 0.001. Furthermore, the interaction between phoneme type and proficiency level was significant in the analysis by subjects $F1$ (1, 84) = 11.76, MSE = 0.26, p = 0.001, and in the analysis by items $F2$ (1, 118) = 4.46, MSE = 0.348, p = 0.037. The calculated effect size (Partial Eta Squared) was large for phoneme type, proficiency level and the interaction between proficiency and phoneme type.

As reported earlier, phoneme type had a significant impact on reaction times in general. However, the size of this effect varied across the proficiency groups as indicated by the significant interaction between phoneme type and proficiency level. As shown in Figure 23, the average reaction times for the monolingual critical condition was slower than for the control condition across the two proficiency groups. In addition, the reaction times for the bilingual critical condition were slower than for the control condition.

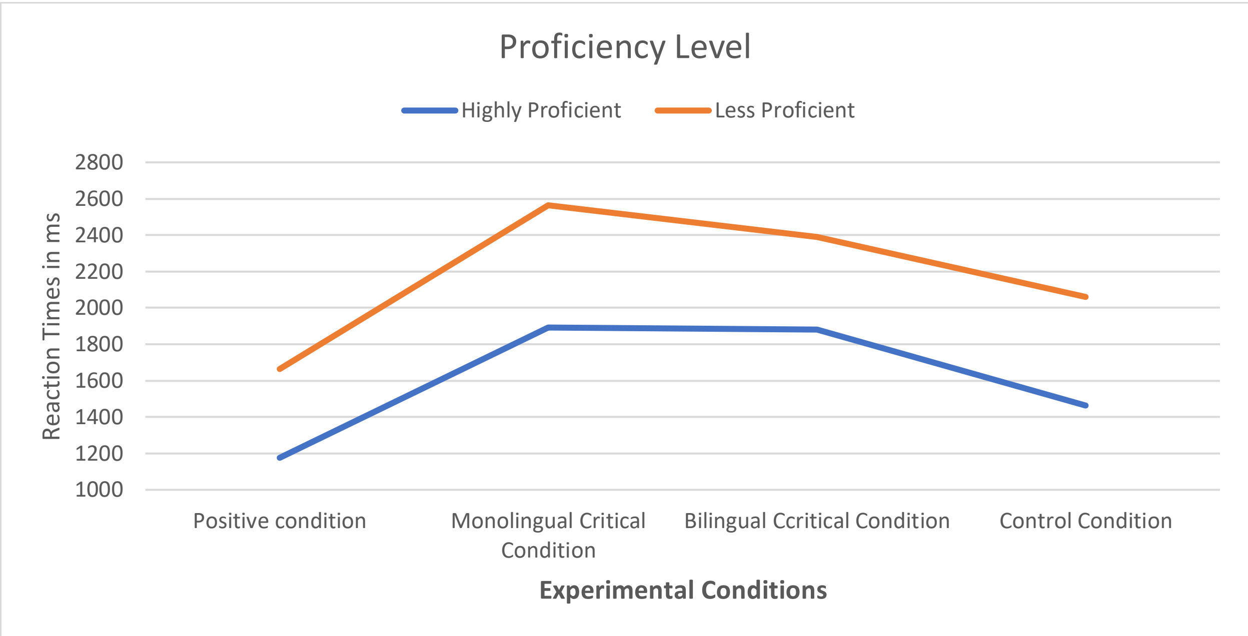


Figure. 23 The effects of phoneme type on reaction times across the highly and less proficiency.

In contrast, for the positive condition, faster reaction times were observed compared to the control and the monolingual and bilingual critical conditions. Moreover, the difference between the two critical conditions was smaller for the highly proficient bilinguals; whereas, for the less proficient bilinguals the difference between the two critical conditions was larger. To examine the significance of the differences noted in reaction times between the four experimental conditions, a paired t-test was conducted. The results showed that the difference in reaction times between the monolingual critical condition and the bilingual critical condition was statistically insignificant (i.e., p > 0.001) for the highly proficient bilinguals, and significant (i.e., p < 0.001) for the less proficient bilinguals. The results are reported in Table 51.

Proficiency Level	Group 1	Group 2	Average reaction times (SD)		Difference between group 1 & group 2 (SE)	t-statistics (df)	p-val of t-test (Bonferroni correction)	Effect size (Cohen's d)	Effect size tabulation
			Group 1	Group 2					
Highly proficient bilinguals	Monolingual critical	Control condition	1892 (59)	1463 (56)	428 (39)	10.76 (39)	< 0.001	1.167	Large
	Bilingual critical	Control condition	1882 (55)	1463 (56)	418 (54)	7.68 (39)	< 0.001	1.173	Large
	Positive condition	Control condition	1176 (48)	1463 (56)	- 287 (33)	- 8.61 (39)	< 0.001	0.865	Large
	Monolingual critical	Bilingual critical	1892 (59)	1882 (55)	10.3 (49)	0.21 (39)	= 0.835	0.028	small
	Monolingual critical	Positive condition	1892 (59)	1176 (48)	716 (44)	16.25 (39)	< 0.001	2.103	Large
	Bilingual critical	Positive condition	1882 (55)	1176 (48)	705 (49)	14.18 (39)	< 0.001	2.140	Large
Less proficient bilinguals	Monolingual critical	Control condition	2564 (37)	2061 (36)	502 (24)	20.92 (45)	< 0.001	2.007	Large
	Bilingual critical	Control condition	2388 (31)	2061 (36)	327 (25)	12.67 (45)	< 0.001	1.429	Large
	Positive condition	Control condition	1664 (32)	2061 (36)	- 396 (26)	- 14.82 (45)	< 0.001	1.705	Large
	Monolingual critical	Bilingual critical	2564 (37)	2388 (31)	175 (25)	6.84 (45)	< 0.001	0.744	Large
	Monolingual critical	Positive condition	2564 (37)	1664 (32)	899 (29)	30.51 (45)	< 0.001	3.766	Large
	Bilingual critical	Positive condition	2388 (31)	1664 (32)	724 (33)	21.75 (45)	< 0.001	3.328	Large

Table. 51 The results of the paired t-test to examine the differences in reaction times between the four experimental conditions.

As detailed in Table 51, the differences between reaction times in the monolingual critical condition and the control condition, and between the bilingual critical condition and the control condition were statistically significant (reaction times for both critical conditions were larger compared to that for control condition) for both the highly and less proficient bilinguals, as all the corresponding p-values were < 0.001. The difference in reaction times between the positive condition and the control condition was statistically significant, as was that between the other critical conditions and the positive condition for all bilinguals regardless of their proficiency level (all p values were < 0.001, as shown in Table 51. Considerably faster reaction times were reported for the positive condition compared to the other experimental conditions for all participants. These findings will be discussed in section 8.3.

The mixed ANOVA analysis indicated a significant main effect of proficiency level as a between- subjects factor; thus, we conducted a follow up independent sample t-test to further explore the significance of the differences noted, for all the four experimental conditions. As summarized in Table 52, the average reaction times for the highly proficient bilinguals were significantly faster compared to that for the less proficient bilinguals across all four experimental conditions. The calculated Cohen's d was > 0.8 for all comparisons, indicating a "large" effect size for proficiency in all conditions.

Phoneme Type	Average RT (SE)		The difference between highly & less Proficient bilinguals (SE)	T-statistics (df)	p-value	Effect size (Cohen's d)	Effect size tabulation
	Highly proficient bilinguals	Less proficient bilinguals					
Control condition	1463 (56)	2061 (36)	- 597 (65)	- 8.89 (67.28)	< 0.001	1.00	Large
Monolingual critical condition	1892 (59)	2564 (37)	- 671 (68)	- 9.54 (67.46)	< 0.001	1.07	Large
Bilingual critical condition	1882 (55)	2388 (31)	- 506 (62)	- 7.89 (62.33)	< 0.001	0.90	Large
Positive condition	1176 (48)	1664 (32)	- 488 (58)	- 8.56 (84)	< 0.001	0.93	Large

Table. 52 The results of the independent samples t-test results to examine the differences in reaction times between the highly proficient and the less proficient bilinguals across the four experimental conditions.

8.2.2 Accuracy Analysis (ACC)

A mixed ANOVA was conducted with proficiency level (highly and less) as the between-subjects factor, and phoneme type (positive, monolingual critical, bilingual critical and control) as the within-subjects factor.

The results of the mixed ANOVA revealed a significant main effect of phoneme type in the analysis by subjects $F1(1, 84) = 4.427$, $MSE = 160.65$, $p = 0.038$, and in the analysis by items $F2(1, 118) = 95.38$, $MSE = 7.45$, $p < 0.001$. Additionally, the main effect of proficiency level was significant in the analysis by subjects $F1(1, 84) = 25.35$, $MSE = 5304.74$, $p < 0.001$, and in the analysis by items $F2(1, 118) = 44.54$, $MSE = 6782. 92$, $p < 0.001$. Furthermore, the mixed ANOVA indicated no significant interaction between phoneme type and proficiency level in the analysis by subjects $F1(1, 84) = 0.002$, $MSE = 0.079$, $p = 0.963$, and in the analysis by items $F2(1, 118) = 0.026$, $MSE = 2.23$, $p = 0.872$. As shown in Figure 24, the effect of phoneme type on accuracy rate was the same for both proficiency groups.

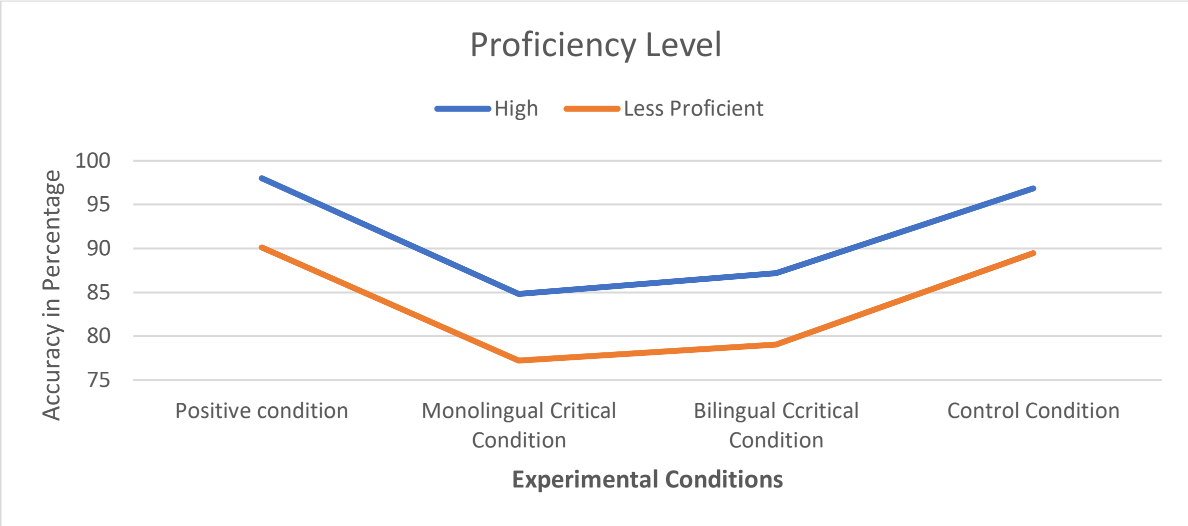


Figure. 24 The effect of phoneme type on the performance of less and highly proficient bilinguals in terms of accuracy rate.

What stands out in Figure 24 was that in the monolingual critical condition, the accuracy rate was considerably low compared to the control condition for both highly and less proficient bilinguals. In addition, the accuracy rate in the bilingual critical condition was low compared to the accuracy rate in the control condition for both proficiency groups. However, when comparing the two critical conditions, it was apparent that the monolingual and bilingual critical phonemes reduced the accuracy rate for both highly and less proficient bilinguals equally. However, the phonemes in the positive condition produced a higher accuracy rate relative to the control condition.

To investigate the significance of the above noted differences, a follow up paired t-test was performed. The results are reported in Table 53.

Proficiency level	Group 1	Group 2	Average ACC (SD)		Difference between group 1 & group 2 (SE)	T-statistics (df)	p-val of t-test (Bonferroni correction)	Effect size (Cohen's d)	Effect size tabulation
			Group 1	Group 2					
Highly proficient bilinguals	Monolingual critical	Control condition	84.94 (14.9)	96.83 (6.03)	- 12.21 (1.96)	- 6.22 (37)	< 0.001	1.171	Large
	Bilingual critical	Control condition	87.22 (10.7)	96.83 (6.03)	- 9.54 (1.39)	- 6.84 (38)	< 0.001	1.132	Large
	Positive condition	Control condition	98.00 (3.4)	96.83 (6.03)	1.17 (0.92)	1.27 (39)	0.426	0.246	Small
	Monolingual critical	Positive condition	84.94 (14.9)	98.00 (3.4)	- 13.08 (2.34)	- 5.59 (37)	< 0.001	1.422	Large
	Bilingual critical	Positive condition	87.22 (10.7)	98.00 (3.4)	- 10.73 (1.66)	- 6.45 (38)	< 0.001	1.509	Large
	Monolingual critical	Bilingual critical	84.84 (14.9)	87.22 (10.7)	- 2.86 (1.90)	- 1.50 (36)	0.283	0.228	Small
Less proficient bilinguals	Monolingual critical	Control condition	77.21 (9.7)	89.49 (9.4)	- 12.09 (1.67)	- 7.22 (40)	< 0.001	1.240	Large
	Bilingual critical	Control condition	79.03 (10.7)	89.49 (9.4)	- 9.79 (1.41)	- 6.94 (43)	< 0.001	0.995	Large
	Positive condition	Control condition	90.11 (7.3)	89.49 (9.4)	0.85 (1.40)	0.60 (44)	> 0.999	0.101	Small
	Monolingual critical	Positive condition	77.21 (9.7)	90.11 (7.3)	- 13.01 (1.74)	- 7.47 (41)	< 0.001	1.511	Large
	Bilingual critical	Positive condition	79.03 (10.7)	90.11 (7.3)	- 11.01 (1.54)	- 7.14 (44)	< 0.001	1.215	Large
	Monolingual critical	Bilingual critical	77.21 (9.7)	79.03 (10.7)	- 2.10 (1.92)	- 1.09 (40)	0.561	0.202	Small

Table. 53 Paired t-test results to test the effect of conditions on accuracy across different proficiency subgroups.

As detailed in Table 53, the differences between both critical conditions and the control condition were statistically significant ($p < 0.001$) for both the highly and less proficient bilinguals. The average accuracy rate for the critical conditions were lower when compared to the control condition for both groups. In contrast, the difference between the positive condition and the control condition was insignificant for the highly proficient group ($p = 0.426$) and the less proficient group ($p > 0.999$).

In addition, the difference between the monolingual critical and the bilingual critical condition was insignificant for the highly proficient group ($p = 0.283$), and for the less proficient group ($p = 0.561$). Finally, the increase in accuracy rate that was apparent in the positive condition was significantly higher than in the control condition and the other critical conditions, and all the p values were < 0.001 for the less and highly proficient groups.

Since there was a main effect of proficiency level, an independent sample t-test was conducted to investigate the differences between the two proficiency groups. As shown in Table 54, the highly proficient bilinguals were more accurate than the less proficient groups.

Experimental conditions	Average ACC (SD)		Differences between the High & less proficient groups (SE)	T-statistics (df)	p-val	Effect size (Cohen's d)	Effect size tabulation
	Highly Proficient Bilinguals	Less Proficient Bilinguals					
Control condition	96.83 (6.3)	89.49 (9.4)	7.34 (1.70)	4.31 (75.66)	< 0.001	0.46	Middle
Monolingual critical	84.81 (14.9)	77.21 (9.7)	7.38 (2.85)	2.59 (62.82)	0.012	0.29	Small
Bilingual critical	87.21 (10.7)	79.03 (10.7)	8.18 (2.35)	3.48 (82.00)	< 0.001	0.28	Small
Positive condition	98.00 (3.4)	90.11 (7.3)	7.89 (1.21)	6.53 (65.91)	< 0.001	0.68	Middle

Table. 54 The results of the independent samples t-test to test differences in accuracy between the highly and less proficient bilinguals across all experimental conditions.

As noted in Table 54, the differences between the highly and less proficient bilinguals were statistically significant under all conditions, as all the p values were $< .05$. The calculated Cohen’s d indicated a moderate effect size for proficiency in the control and positive conditions; whereas, for the critical conditions a small effect size was found. We now discuss the implications of these findings.

8.3 Discussion

a. Are the L1 distinct phonemes (i.e., non-target phonological segments) activated at the phonological level? And what is the impact of this activation?

Similar to experiment four, the performance of the Arabic-English bilinguals in the four experimental conditions varied as a result of our manipulation of the phoneme type. The statistical comparisons conducted confirmed that all the participants responded faster when the phoneme was part of the picture name in English; i.e., in the positive condition compared to the control condition (a significant difference of 342 ms). The most interesting finding from this experiment was that, across all the participants, distinct phonemes yielded slower reaction times relative to the control condition (a significant difference of 465 ms), and relative to the shared phonemes (a difference of 92 ms). In addition, similar to experiment four, the shared phoneme condition produced slower reaction times compared to the control condition. These findings confirmed the hypothesis that the target and non-target phonological representations are activated at the phonological level regardless of the type of L1 phoneme used. In other words, the L1 shared and distinct phonemes are all active at the phonological level and impede the performance of participants in this monolingual task. It was hypothesized that the participants would exploit the distinct phonological and orthographical features of the L1, and thus readily form a negative response upon encountering distinct phonemes. On the contrary, these phonemes were processed as candidates during the decision-making process, as was apparent from the slower reaction times obtained in that condition. Moreover, the accuracy rate fell significantly in the case of shared and distinct phonemes relative to the control condition (an estimated difference of 12% and 10% respectively) across all the participants. These findings indicated that the participants were more likely to respond positively in the shared and distinct phonemes trials, due to the automatic activation of the L1 non-target nodes at the phonological level. The fact that the participants provided a positive answer to the L1 distinct phoneme when they were expected to reject it instantly prior to the presentation of the picture suggests that the two lexical nodes were simultaneously activated upon the presentation of the picture at the phonological level, and both were considered when responding to the task in hand. Therefore, at the phonological level we have evidence that not only were the shared phonemes activated, but also that language distinct phonemes were similarly active in different script bilinguals and that differences associated with script did not inhibit the activation of the non-target segments.

The findings of this experiment replicated the results found in experiment four (chapter seven), and confirmed the view of the cascaded activation of non-target nodes at the phonological level with different script bilinguals (Korean-English: Moon and Jiang 2012), in line with same script bilingual findings (Catalan-Spanish: Colomé 2001). Moreover, the current results extend published findings (e.g., Dell 1986; Humphreys et al. 1988) that suggest cascaded activation is not limited to shared phonemes across the bilinguals' two languages (i.e., Arabic-English speakers), but also holds true for phonemes that are distinct in the L1. Of course, further investigations are needed for same and different script languages before making generalizations. To the best of the author's knowledge, this is the first attempt to investigate the question of whether cascaded activation is applied to L1 distinct phonemes while processing in L2 using a phoneme monitoring task. It remains unclear whether the same principle is applied to L2 distinct phonemes when the task is performed in L1.

b- What is the effect of proficiency level on the cross-language phonological activation?

Regarding the effect of proficiency level, the results of the reaction times and accuracy analysis showed that the highly proficient bilinguals were faster and more accurate than the less proficient bilinguals across all conditions. Moreover, both groups were similarly affected by the type of phoneme presented, and the effect size of the critical conditions proved to be large for both. However, the less proficient group were significantly more affected by the distinct phonemes than the shared phonemes. In the following paragraph, we discuss these two major findings, along with their implications.

The finding that the highly proficient bilinguals performed better than the less proficient bilinguals, echoes that reported in experiment four (phoneme monitoring task chapter seven), and the first three experiments that employed the masked primed picture naming task (chapters four, five, & six). This was explained in relation to the weak links theory (Kroll and Stewart 1994). As discussed in chapter two, it can be assumed that the less proficient bilinguals have delayed L2 lexical access due to weak links between concepts and the L2 lexicon compared to the highly proficient bilinguals, and that the lexical links between L2 and L1 are strong due to frequent access (i.e., the access to concept of L2 words is mediated through L1 words). Kroll et al. (2010) argued that for the less proficient bilinguals access form L2 words to concept may be easily accomplished, whereas access from the concept to the L2 words is effortful, especially in production tasks [17].

Therefore, to achieve the task requirement, the less proficient bilinguals in experiment five had to: (i) activate the conceptual representation of the target picture; (ii) retrieve the corresponding L2 lexical node; (iii) phonologically encode it; (iv) match the phoneme presented across the target phonemes; and (vi) provide a decision. It was argued here that the difficulty lies in the L2 lexical retrieval stage, due to weak links from concept to L2 and strong lexical links between L2 and L1(which makes the L2 node more vulnerable to the L1 interference effect).

[17] Kroll et al. (2010) argued that in word recognition tasks the less proficient learners were able to directly access the meaning of L2 words and this is due to a strength asymmetry of the link between L2 and concept. Links from L2 to concept are not as weak as those from concept to L2 words.

Moreover, the within proficiency group comparisons revealed that shared and distinct phonemes had a similar size of effect on the performance of highly proficient bilinguals (i.e., an insignificant difference in RT between the two critical conditions of 10 ms). Meanwhile, the distinct phonemes had a greater effect than the shared phonemes on the performance of less proficient group (a significant difference in RT between the two critical conditions of 180 ms). A possible explanation for this considerable drop in the performance of less proficient bilinguals in that condition could be attributed to difficulties recognising the phonetic symbols for the distinct phonemes. We argue that they might not be recognized as fast as the shared phonemes by the less proficient bilinguals. In the preparation stage, we asked the participants to study the phonetic symbols and we conducted a practice session followed by a testing exercise to ensure accurate recognition had been achieved. However, we did not check how fast they were able to respond; i.e., they might need more practice to hasten the speed of recognition of the distinct phonemes. Therefore, this might be a potential cause for the slower reaction times.

Another possible explanation was related to strength differences of the two proficiency groups in the language control mechanism. The literature widely reports a corelation between proficiency level and language control mechanism; i.e., as proficiency increases, the language control mechanism is enhanced during language processing (e.g., Costa and Santesteban 2004; Costa, Santesteban, et al. 2006; Blumenfeld and Marian 2007). Studies involving language switching tasks (e.g., Costa and Santesteban 2004; Costa, Santesteban, et al. 2006) also reported that the less proficient bilinguals applied a stronger inhibition of L1 representations (dominant language) compared to L2; thus, switching into the L1 was effortful (see section 2.3.4). This was mainly because more time was needed to overcome their inhibition. These studies concluded that highly proficient bilinguals, who are dominant in L1 and L2 can apply a strong inhibiting effect over the two languages; whereas, less proficient bilinguals strongly inhibit the L1 relative to the L2.

Examining the results of experiment five, it appears that for the less proficient bilinguals greater control was required over the L1 names, which were triggered by distinct phonemes compared to shared phonemes. To validate this hypothesis, let us look at the underlying processes involved in this task. As discussed previously in this section, the results indicated that when a picture was presented, its names in L1 and L2 were activated, regardless of the type of phoneme presented ahead. When responding to the task demands, the participants had to supress/ignore the activation of the non-target lexical node, focus on the target lexical node, and determine whether the phoneme was a part of it or not. Presenting an L1 phoneme did not make the process any easier, as it added additional support to the activation of the L1 name that participants had to ignore/supress. However, the results showed that presenting an L1 distinct phoneme produced a slower reaction times compared to the L1 shared phoneme. We cannot conclude that words with distinct phonemes are harder to supress or ignore, because all the L1 names of pictures in this experiment had a distinct phoneme. Therefore, the mere presence of a distinct phoneme in the L1 picture name does not account for this variation in reaction times. The only possible cause appears to be the presentation of a distinct phoneme ahead of the picture. It could be that the distinct phonemes had a greater effect in terms of their activation of the L1 name compared to the shared phonemes. Therefore, the L1 picture names that received additional activation form the distinct phonemes were harder to supress, and this interference was reflected in the slower reaction times observed in that particular condition. According to Green (1998), the resolution of cross lexical competition is dependent on the relative activation level of lexical competitors, so the higher the activation of the two lexical representations, the longer it will take to suppress the non-target lexical nodes to allow selection of the target lexical representation.

Overall, the findings of experiment five suggest that the two languages were activated simultaneously, and that this activation extended to the phonological level. Moreover, proficiency level did not modulate the manner of language selection and lexical access or the flow of activation.

8.4 Conclusion

This experiment investigated lexical access for Arabic-English bilinguals. It also explored whether proficiency level modulates cross language phonological activation of lexical selection for different script bilinguals. We replicated experiment four (Chapter seven) to determine whether the same pattern of findings would be obtained using L1 distinct phonemes. The results indicated the existence of parallel activation of the two languages at the phonological level, regardless of the type of L1 phoneme used and any script differences. Additionally, the results showed that proficiency level did not modulate the cross-language phonological activation; as the performance of the bilinguals in the two groups was significantly affected by the presence of the distinct and shared L1 phonemes. Based on these findings, we conclude that the manner of lexical access in different script bilinguals is non-selective, being similar to that of same script bilinguals. Moreover, that the flow of activation cascaded to the phonological level. We will now summarise and critically evaluate the experiments and findings detailed here, reflecting briefly on the limitations and recommendations for future work identified.

Chapter 9 General Discussion

9.1 Summary of the major findings

For many years, psycholinguistic studies have investigated lexical access in the context of bilinguals' word production. These studies, however, were conducted with same script bilinguals, and thus provided theoretical models for cross-language activation in bilinguals, such that the manner of language selection, and the flow of activation was limited to these populations. The main objectives, in the present study, involved examining cross language activation, the manner of lexical/phonological selection, and the flow of activation for different script adult bilinguals (Arabic- English). The study also sought to investigate whether proficiency level (highly and less proficient) would modulate cross-language activation among different script bilinguals. Through this we hoped to confirm or challenge current models of bilingual lexical access, and this was achieved through an innovative methodological approach pertaining to the masked priming paradigm.

The innovative work of examining the process of single word production in different script bilinguals described here is important for two reasons. Firstly, the existing theories may not necessarily describe the process of word production in different-script bilinguals, and so need to be challenged to ensure their relevance to all typical adult populations; and secondly, examining different script bilinguals allowed us to investigate the cross-language phonological activation with the zero orthographic overlap. The experiments were designed to tap into each stage of lexical access and draw informative conclusions about the constantly debated issue of lexical selection (language specific vs non-specific) and the flow of activation (cascaded vs discrete). Five experiments were conducted involving Arabic-English bilinguals. Different groups of Arabic-English participants were recruited for every task. The experiments were conducted consecutively and triggered by early findings: (i) experiment one (chapter four) was conducted to examine the cognate facilitation effect and the identity effect; (ii) experiment two (chapter five) was conducted to replicate the semantic interference effect observed in picture-word interference tasks, and to examine whether the manner of lexical selection is language specific as indicated by the findings reported in experiment one; (iii) experiment three (chapter six) was conducted to identify the locus of the semantic effect found in experiment two (chapter five); (iv) experiment four (chapter seven) was conducted to examine cross-language phonological activation using phonemes shared by the Arabic and English languages; and (v) experiment five (chapter eight) was conducted to establish whether the results obtained in experiment four (chapter seven) could be replicated using L1 distinct phonemes. The following sections summarize the main findings and discuss the implications of these findings for models of lexical access in bilinguals.

In experiment one (chapter four), highly and less proficient Arabic-English bilinguals were asked to name cognate and non-cognate pictures in L2, preceded by a masked prime word in L1. The prime words were either the name of the target picture in the L1 or an unrelated name in L1. The reason for using prime words in L1 was to increase the level of activation of non-target lexical nodes in the non-target language. Thus, if selection by competition exists, it should manifest clearly. More specifically, it was anticipated that naming cognate pictures preceded by phonologically similar L1 names should induce an interference effect if language selection is sensitive to the activation level for both target and non-target lexical nodes. The cognate name in L1 receives activation, not only from the shared conceptual representation, but also from the prime word, thus increasing the level of activation of its corresponding shared and different phonological segments at the phonological level. Similarly, it was anticipated that naming a non-cognate picture preceded by its L1 name should induce an interference effect if lexical selection considers the level of activation of both the target and non-target lexical node. The results showed that priming cognate pictures with their L1 names induced a facilitation effect and enhanced accuracy rates. The findings also supported existing research (Janseen 1999; Hoshino 2006), as participants named cognate pictures more rapidly than non-cognate pictures. The cognate facilitation effect found in the related condition (i.e., the cognate name of the picture in L1) suggests that, although the non-target lexical node and its phonological segments were highly active, they did not interfere with the selection of the target phonological segments. There are two possible explanations for this finding; the first is that the phonological selection only considered activated target phonological segments and thus the manner of selection is language specific, which supports Costa, Caramazza et al.'s (2000) view; the second explanation is that the selection process considers both target and non-target segments (i.e., language non-specific selection), but as the number of non-shared phonemes in the non-target word is few (compared to shared), they were unable to reach the threshold for interference, and this then failed to impact the selection process. However, this alternative explanation is subject to further investigation, as discussed in chapter four (section 4.3). The second major finding was that non-cognate pictures preceded by their names in L1 induced a facilitation effect. This indicates that both target and non-target lexical nodes were active, but lexical selection considers only the target lexical nodes in the target language. The observed facilitation effect took place at the lexical level; i.e., it was attributed to the non-target lexical node, sending additional activation to the target lexical node. The third finding was that script related differences did not modulate lexical access and the manner of lexical/phonological activation. Despite being presented in the L1, related prime words were processed by bilinguals and facilitated the selection of the target lexical node; i.e., script differences did not act as a language cue to inhibit the activation of non-target nodes. In addition, despite the lack of an orthographic overlap, the cognate effect manifests in the production of different script bilinguals. The last major finding was that the same pattern of results emerged regarding naming cognates and non-cognates, and this was observed for both highly and less proficient bilinguals. This suggests that proficiency level did not modulate cross-language activation or the flow of activation. In addition, it was noted that proficiency level did not influence the size of the cognate effect. The only difference to emerge between the two groups was that the highly proficient bilinguals named the pictures faster and more accurately than the less proficient bilinguals.

In experiment two (chapter five), a different group of highly and less proficient Arabic-English speakers were asked to name non-cognate pictures, which were preceded by semantically related masked-prime words and unrelated masked prime words. This experiment assessed whether the reported findings of a semantic interference effect (e.g., Hermans et al. 1998) could be replicated using a different paradigm i.e., the masked priming paradigm in the prime words are not visually available for conscious processing, unlike the distractor words in picture word interference tasks. The presentation time for prime words were manipulated (i.e., 100 ms, 75 ms, and 50 ms) to determine whether the effect would manifest at 100 ms, similar to monolinguals (Alario et al. 2000; Bajo et al. 2003) or whether it could surface earlier. The results revealed that the semantically related prime words presented for 100 ms and 75 ms induced an interference effect when preceding the non-cognate pictures, which is in accordance with the findings of the previous studies that used the picture word interference task. Additionally, these results confirm and extend the findings of the monolinguals studies in that the semantic interference effect was obtained not only at 100 ms but also at 75 ms. The same pattern of findings was found for the less and highly proficient bilinguals. Both proficiency groups experienced the semantic interference effect at 100 and 75 ms, but not at 50 ms, which confirms a similar processing time for the effect to manifest across the two groups. However, the highly proficient bilinguals outperformed the less proficient bilinguals in this task, as with experiment one. In addition, the finding of semantic interference effect in the performance of different script bilinguals indicates that perceived script differences did not inhibit cross-language activation/competition, which concurs with Boukadi et al.'s (2015) findings. Overall, the findings indicate that the non-target lexical node was active, and impacted the selection process at the lexical level and that proficiency level did not modulate the cross-language competition.

In experiment three (chapter six), we compared the performance of the Arabic-English bilinguals (highly and less proficient) in two different tasks; namely the masked priming in the picture naming task and the animacy decision task. The former involves semantic and lexical activation; whereas the latter involves semantic activation only. The objective was to determine the locus of the semantic interference effect. Thus, semantically related masked prime words were presented before the pictures in the picture naming task, and the participants had to decide whether each picture was of an animate or inanimate object in the animacy decision task. If semantic interference effect arises at the conceptual level, then longer reaction times should be found for both tasks when the semantically related prime words preceded the target pictures. However, if the semantic interference effect is located at the lexical level, then a longer reaction times should be found only in the picture naming task, which required lexical activation, in addition to semantic activation. The results showed that semantically related prime words induced interference in the picture naming task whereas in the animacy decision task it induced a facilitation effect. The same pattern of effect was found for both highly proficient and less proficient bilinguals. The only difference was that the size of the semantic effect was greater for the highly proficient bilinguals when compared to the less proficient bilinguals over the two tasks. The findings indicated that the semantic interference effect is not located at the conceptual level, and that this effect might originate at the lexical level. Having several related concepts activated at the conceptual level prompted activation of the corresponding lexical nodes at the lexical level, thereby broadening the selection pool. The selection mechanism had to consider several activated lexical nodes, which competed for selection and thus delayed the naming process in the picture naming task. The findings of a semantic facilitation effect in the animacy decision task suggest that related concepts send activation to the target concept, and thus facilitate the selection of the target concept. In terms of proficiency level, the results indicate that it did not modulate cross-language activation. However, finding a greater sized semantic effect (facilitation and interference) for highly proficient bilinguals suggests that the highly proficient bilinguals were more sensitive to the manipulations of stimuli type in both tasks relative to the less proficient bilinguals. Regarding the role of script differences, we replicated the findings reported in experiments one and two that suggest script differences did not inhibit the activation of the non-target language in a task that requires only conceptual processing (the animacy decision task), and in the other task that requires conceptual and lexical processing (the masked primed picture naming task).

Experiment four (chapter seven) investigated whether the flow of activation was cascaded or discrete, in different script bilinguals. The highly and less proficient participants were presented with phonemes followed by a picture, and they were expected to decide whether the phoneme was part of the L2 name of the picture or not. Three experimental conditions were created: a critical condition in which the phoneme was part of the L1 name of the picture, a positive condition in which the phoneme was part of the L2 name of the picture, and finally a control condition in which the phoneme was not considered part of the picture name in either the L1 or the L2. If the non-target phonological segments corresponding to the activated non-target lexical nodes were active along with the target phonological segments, then it was anticipated that longer reaction times and a low rate of accuracy would be observed in the critical conditions. The results showed that phonemes that were part of the L1 name of the pictures induced longer reaction times, and lower accuracy rates relative to the phonemes in the positive and control conditions. The same pattern of findings was obtained for the less and highly proficient bilinguals. These findings clarify that the activated non-target lexical nodes send activation to corresponding phonological segments at the phonological level. With regard to the effect of proficiency level, the results suggested that had no effect in terms of modulating the flow of activation. In terms of the role of different script, the findings from experiment four suggest that phonological co-activation is automatic in different script bilinguals, even when there is no phonological/orthographical overlap between the two alternatives (i.e., the target and non-target name). Everything considered, this study provides evidence to support the cascaded view of activation, which holds true for both highly and less proficient Arabic-English bilinguals.

Finally, experiment five (chapter eight) tested whether the same pattern of cascaded activation would be obtained if L1 specific phonemes were used in the phoneme monitoring task. As explained in chapter eight, this is mainly because shared and distinct phonemes from the two languages have different type of interactions (see Flege 1987). Furthermore, in an L2 production task for bilingual speakers, shared phonemes across the two languages proved to be more active than L1 specific phonemes, and thus further investigation was needed before making any generalization. Therefore, two groups of Arabic-English bilinguals (highly and less proficient) were asked to decide whether or not the phoneme presented was part of the L2 picture name. There were four types of phonemes: a phoneme that is part of the L2 name of the picture in the positive condition, a language specific phoneme that is part of the L1 name of the picture in the monolingual critical condition 1, a shared phoneme (across the two languages) that is part of the L1 name of the picture in the bilingual critical condition 2, and finally a phoneme that is not part of the picture name in either the L1 or L2 in the control condition. If the non-target lexical nodes activated their corresponding phonological segments, irrespective of the type of phoneme, then longer reaction times and low rate of accuracy should be apparent in the critical condition 1 and 2. However, if only shared phonemes were active at the phonological level, then longer naming latencies and low rate of accuracy should be noticed in only the critical condition 2. The results showed that shared and language specific phonemes induced longer reaction times in critical conditions 1 and 2. Moreover, the same pattern of findings was evident for the less and highly proficient bilinguals. These findings indicate that non-target phonological segments were active at the phonological level, regardless of their type i.e., being shared or language distinct phonemes. Once again, the findings indicate that script differences did not inhibit the activation of the non-target language, nor did they modulate the cascaded flow of activation.

In summary, this study investigated the cross-language activation across different script bilinguals and the manner of lexical selection during word production by applying three different experimental paradigms. It also evaluated the role of proficiency level to establish whether it affects cross-language activation and manner and of lexical selection, by comparing the performance of highly and less proficient bilingual speakers. We will now go on to evaluate the implications of these findings.

9.2 Implications for Models of Lexical Access in Bilingual Production

9.2.1 Is lexical access non-selective in different script bilinguals?

Former research which investigated the word production mechanism in bilinguals whose languages share the same Roman alphabet concurs with non-selective activation (Hermans et al. 1998; Costa, Colomé, et al. 2000; Colomé 2001; Gollan and Kroll 2001; Costa et al. 2003; Costa and Santesteban 2004). The present research examined the performance of different script bilinguals (Arabic-English speakers) over five different experiments. The overall findings corresponded with previous studies regarding the co-activation of the two languages at the lexical level. The Arabic- English bilinguals in this study could not shut down the non-target language completely; i.e., the language that is not required in the task.

The activated conceptual representations automatically spread to the lexical nodes in the two languages at the lexical level, regardless of the task demand. This pattern was evident in all the tasks employed in this study: a) priming a cognate and a non-cognate picture with a masked translation prime word in the non-target language in the picture naming task, b) priming a non-cognate picture with a semantically related masked prime word in the non-target language in the picture naming task, and finally in c) the phoneme monitoring task with shared phonemes and with language specific phonemes. In these tasks, the Arabic script (i.e., in L1 prime words) did not act as a language cue to limit activation to the target lexical nodes only. The results of this study are compelling enough to conclude that script differences did not modulate cross-language activation. Later in this chapter, the issue of whether parallel activation extends to the phonological level or not will be discussed thoroughly.

As reported in Chapter one, the evidence for co-activation in single word production comes from various experimental paradigms, such as, the phoneme monitoring task (e.g., Colomé 2001), the cognate picture naming task (e.g., Costa, Caramazza, et al. 2000), language mixing/switching (e.g., Kroll et al. 2008), and the picture word interference task (e.g., Hermans et al. 1998). The majority of these studies, aside from the phoneme monitoring task, investigated cross-language co-activation in the presence of the stimuli in the non-target language, which may have had an effect on the activation states in both languages (Grosjean 2001; Kroll et al. 2006). Here, we presented new supporting evidence for co-activation across the two languages from the masked priming paradigm in the picture naming task, which has the advantage of unconscious processing of the L1 masked stimuli. This study also extends the findings regarding the non-selective lexical access among same script bilinguals to include different script bilinguals. Finally, the findings of the current study correlate with recent research conducted on different script bilinguals with regard to non-selective lexical access (e.g., Boukadi et al. 2015). These recent studies were aimed at replicating the findings of same script studies. Thus, the investigations were performed using similar traditional methods, but with different groups of bilinguals, such as the Korean-English bilinguals in a phoneme monitoring task (Moon and Jiang 2012), Japanese-English bilinguals in a picture naming task (Hoshino 2006) , and Arabic-French bilinguals in the picture word interference task (Boukadi et al. 2015). In short, the importance of this study's finding is that it confirms and extends previous findings concerning the co-activation of the lexical nodes across the two languages.

9.2.2 Is lexical selection language specific or non-specific?

The nature of the lexical selection process has been heavily debated (e.g., De Bot 1992; Green 1998; Roelofs 1998). The findings from several studies favoured the view of non-selective lexical selection; however, other results favoured the view of selective lexical access. In the present research, the results of naming cognates and non-cognates which were preceded by the L1 translation name in the masked priming paradigm supported language specific selection. However, the results of naming non-cognates that were preceded by semantically related prime words, indicated a language non-specific selection that entails competition.

Furthermore, the findings of the last two phoneme monitoring tasks indicated the activation of the non-target phonological representations at the phonological level, which impeded participants' decisions. Although the main objective of the two phoneme monitoring tasks was to investigate whether or not the flow of activation cascades to the phonological level, the results were informative in terms of the issue of lexical selection. Thus, reference to these findings will be made when discussing the lexical selection process along the following lines.

Overall, the results indicate that the lexical selection mechanism is not entirely language specific or non-specific. At this point, a question arises concerning what governs this selection process. Here, we propose a model in which the selection process complies with the preverbal message (i.e., the target picture and the task demand to provide an L2 name) but is still open to any available linguistic cues; i.e., it will process the linguistic information coming from the distractor/prime word, and consider all highly active nodes and verify them against the preverbal message. For example, in the picture word interference task, the selection process selects the target lexical node but notices the presence of a highly active non-target node (i.e., corresponding to distractor/prime), which it starts to verify it against the preverbal message. The duration of this verification process is dependent on the type of linguistic information the distractor/ prime words share with the target node; specifically, whether there is partial or complete semantic overlap. Complete semantic overlap would be expected to facilitate the selection process; however, partial overlap would lead to interference. In the following lines, we explain this view in greater depth.

In the three masked priming paradigm experiments, the manipulation of prime type led to different outcomes; i.e., semantically related prime word induced interference; whereas, the translation prime word induced facilitation. The preverbal messages direct the selection mechanism when choosing a lexical node that corresponds to the target picture in the target language, and selection is achieved by choosing the lexical node that has the highest activation level. In the identity condition, distractor/prime words induced the activation of the non-target name, which had already been simultaneously activated as per the compelling evidence of the co-activation reported earlier. Therefore, the pool of selection is much the same (i.e., the two alternative lexical nodes in the two languages). The selection mechanism observes the highly active alternative and rejects it after verifying it against the preverbal message, which is to include L2 lexical nodes only. The verification process is rapid, due to the limited number of activated alternatives compared to the other conditions. The reason for the faster retrieval reported in the identity condition relative to the unrelated condition lies in the number of activated alternatives in each condition. In the unrelated condition, unrelated concepts and lexical nodes are activated alongside the target nodes and their counterparts in the other language.

Conversely, in the semantically related condition, the distractor/prime words induced the activation of related concepts and lexical nodes in the two languages. Thus, the number of activated lexical items at the lexical level was higher compared to the identity condition. The selection mechanism considers all the activated lexical nodes and verifies them against the preverbal message. As the number of activated lexical nodes increases, the speed of the verification process becomes impeded. It is important to note here that the type of relationship between these activated nodes plays a vital role here. It is harder to process and verify related nodes than unrelated nodes. In the related and unrelated conditions, the number of activated nodes were high, but it was only those related nodes that impeded the selection process. The unrelated nodes are not supported at the conceptual level (from the target concept), thus their activation level was not sufficiently high to significantly impede the verification process.

It is crucial to point out that in this model, selection at the lexical level is not final, as the process functions at the phonological level, and reactive to any linguistic cues that it comes across until it reaches the articulatory level. This explains why the different types of phonological distractors induced an interference/facilitation effect at the phonological level. For example, in the picture word interference tasks, a phono-translation distractor induced interference, whereas a phonologically similar distractor word induced a facilitation effect (e.g., Hermans et al. 1998). Whether the selection process at the phonological level is the final stage or not (i.e., the debate of the locus of language selection), will be discussed subsequently in section 9.2.4.

The implication of the present research regarding lexical selection is that it addresses the contradictory outcomes in the literature. It has merged the two classical views of the lexical selection process, and created a model that merges previously published findings. Moreover, it has added to the body of knowledge regarding the manner of lexical selection in different script bilinguals; i.e., comparing them with what we know already about same script bilinguals and monolinguals. We concurred that the different script bilinguals experienced the same cross-language competition/facilitation effect as the same script bilinguals (which is typically reported in the picture word interference tasks with semantically related, phono-translation, translation and phonologically similar distractor words). The Arabic-English bilinguals experienced interference when a semantically related prime word was used, which corresponds to the outcomes of the picture word interference tasks in same script studies (e.g., Hermans et al. 1998) and a recent different script study (Arabic- French bilinguals) (Boukadi et al. 2015). Furthermore, the bilinguals in this study experienced a facilitation effect when a translation prime word (the identity effect) was presented, which replicates the findings of the picture word interference task in same script studies (e.g., Costa and Caramazza 1999) and different script studies (Japanese-English speakers: Hoshino 2006). Therefore, it can be argued that the proposed model for lexical selection not only describes the selection mechanism for different script bilinguals, but also accommodates same script bilinguals due to the symmetrical findings reported earlier.

In summary, we conclude that as the activation flows from the conceptual level to the phonological level irrespective of the selection process, so does the selection process. It functions at the conceptual level all the way up to the phonological level, trying to identify a matching candidate for all the pre-set cues it has received from the preverbal message.

9.2.3 Is the flow of activation cascaded or discrete in different script bilinguals?

With regard to the flow of activation, former research on same script bilinguals produced contradictory evidence. More specifically, the discrete view (e.g., Levelt 1989: and Levelt et al. 1991) assumes that activation flows from the conceptual system to the lexical level, and that only the lexical node selected can activate corresponding phonological segments. Whereas the cascaded view (e.g., Caramazza, 1997; Cutting & Ferreira, 1999) claims activation spreads from the conceptual system to the lexical level, and on to the phonological level. However, there is growing evidence supporting the cascaded view in similar script bilinguals (e.g., Kroll et al. 2000).

In the present research, our investigations into the different script bilinguals revealed a cascaded flow of activation, which indicates that script differences did not act as language cues and inhibit the activation of the non-target lexical nodes at the lexical level, nor did they modulate the activation of the corresponding phonological segments at the phonological level.

In addition, the activation of non-target lexical nodes flows from the lexical level to the phonological level irrespective of the process of lexical selection. As mentioned in the previous section, lexical selection takes place initially at the lexical level, but this does not prevent activation of non-target phonemes at the phonological level. The activated non-target phonemes are capable of interfering or facilitating the selection process if they receive additional phonological activation, for example, from the distractor/prime words presented in the tasks. The results for naming cognate pictures in the masked priming paradigm indicated an increase in the size of the faciliatory effect when the non-target language was presented as a prime word. Moreover, the results of the two phoneme monitoring tasks showed that presenting a non-target phoneme in the non-target language produced an interference effect. All considered, identifying an interference and a facilitation effect, supports the view that non-target lexical phonemes are active at the phonological level. It also suggests there is a selection process that takes into account the phonological cues available at that level, besides those obtained from the preverbal message. This point will be discussed further in the next section.

The finding regarding the cascaded flow of activation is in line with findings in same scripts studies that use phoneme monitoring tasks (e.g., Colomé 2001) and cognate naming tasks (e.g., Costa, Caramazza, et al. 2000). Additionally, it is in accordance with the findings of different script studies, such as those of Moon and Jiang (2012), who tested Korean-English speakers in a phoneme monitoring task, and the findings of Hoshino (2006) who used the cognate naming task when investigating Japanese-English bilinguals. Moreover, the outcomes of this study confirmed the findings of Boukadi et al.'s (2015) research, investigating Arabic-French bilinguals in a picture word interference task in L2. L1 distractors words, which were manipulated phonologically, induced a phono-translation facilitation effect, and a phonological facilitation effect.

In conclusion, we concur with the growing body of evidence supporting cascaded flow arising from same and different script studies investigating the phonological activation of non-target lexical nodes when using cognate naming tasks, phoneme monitoring tasks and picture word interference tasks. The current study contributed to the findings in the literature, with results also confirming a cascaded flow. What is innovative here is that this contribution proceeds from a different method; i.e., the cognate picture naming in masked priming task, as well as in traditional phoneme monitoring task.

9.2.4 How does the selection process take place? Is there a specific locus of selection or multiple points?

As discussed in the previous sections of our proposed model, lexical selection at the lexical level is not final, as the non-target lexical nodes will activate their corresponding phonological representations, which may then either interfere with or facilitate the selection process at the phonological level if their activation level is boosted. One may argue that since there is compelling evidence for the phonological activation of non-target segments, then it is logical to argue that the selection process does not occur at the lexical level, and that there is only one selection process at the phonological level. If this is true, then how can we explain the findings of the semantic interference effect? From our perspective, this effect is clear evidence of a selection process at the lexical level, as it indicates an impeded decision, i.e., the selection mechanism considers semantically related distractors/primes (e.g., in the picture word interference task & in the picture naming in masked priming paradigm) as possible candidates for selection. It is important to note that experiment three attempted to identify the locus of the semantic interference effect, in particular whether it is located at the concept system or at the lexical level, and the findings pointed to the lexical level as the main locus of this effect.

Notably, in their recently published work, Blanco-Elorrieta and Caramazza (2021) recently argued that there is a common selection mechanism at all linguistic levels. The assumptions of their model are:

(i) The linguistic systems in bilingual speakers are fully integrated;

(ii) There is cross-language activation at all linguistic levels (the lexical, morpho-syntactic, and phonological levels);

(iii) That the selection mechanism operates at all linguistic levels and can be affected independently at each level;

(iv) There is no inhibition mechanism.

According to this model, the selection process is governed by one principle, which requires the selection of only the most highly active item. Moreover, it postulates that the activation level of these items is determined by several factors, which are: the frequency of each individual element in each language, the proficiency level of the bilingual speaker, intended semantic meaning, communicative context and temporal effects. What this model and ours share are the assumption of a common selection mechanism at all levels, vulnerability (i.e., selection is not permanent and could change at subsequent levels), and the fact that no inhibition control is needed to perform the selection. However, there is one point of disagreement, which is the principle that guides the selection process. In Blanco-Elorrieta and Caramazza's model, the selection process relies heavily on the level of activation, whereas in our model the selection process is governed by the level of activation and cues pre-set in the preverbal message. Their model does not explain why different effects manifested in the various word production tasks, as reported in the research here. If the selection process directly selects the item that reaches the activation threshold, then no semantic/phonological interference effect should be evident in the production tasks, such as the picture word interference tasks. This interference effect implies that there are two alternatives that are highly active and compete for selection, and this competition is resolved by reviewing the pre-set cues in the preverbal message (i.e., select only the item in the target language). Another important point worth mentioning here is that Blanco-Elorrieta and Caramazza (2021) claim that proficiency level may modulate the activation level of these items, and this could account for differences in the size of the semantic interference effect found between the two proficiency groups in this research. Although the same effect was obtained for both the highly and less proficient bilinguals across all tasks, the size of the interference effect (when priming pictures by a semantically related masked prime in L1 in experiment two) differed for both groups. The less proficient bilinguals produced a larger effect size as compared to the highly proficient group.

Thus far, we have indicated that selection at the lexical level is not permanent, and could be influenced at the subsequent stage, i.e., the phonological level. It is time now to discuss the process of phonological selection; i.e., whether this selection is the final stage before the articulatory stage. The assumptions of our proposed model state that the main criterion for selection is first to consider those representations that have a high activation level, and second to choose items that correspond to the preverbal message. Thus, when the selection process comes across an item with a high level of activation, it will validate it first against the cues pre-set from the preverbal message. At the phonological level, when introducing a phonologically similar distractor/prime word (e.g., cognate L1 words), the selection process can be administered with ease. This is mainly because while the selection process is considering target phonemes that correspond to pre-set cues in the preverbal message, it observes a high level of activation of non-target phonemes that are similar to those target phonemes. This co-activation acts as a confirmation of the selected phonemes; thus, a decision is facilitated. In contrast, when the activation level of the non-target phonological segments is high, it impedes the selection process as too many highly active segments are available for selection (e.g., the phono-translation condition in Hermans et al. 1998). Here, the selection process has to validate those active segments against pre-set cues and start narrowing down options to include only those corresponding to the target language. Once the selection process is complete, the selected phonemes reach the articulatory level and the word can be retrieved. Therefore, we argue that once the phonological selection is made and assuming no further cues or distractors are presented after that stage, the selected item will finally reach the articulatory stage.

9.2.5 Does proficiency level modulate the cross- language activation?

With regard to this question, both bilingual groups in this study experienced cross language activation. The findings suggest the manner of lexical activation was similar for both proficiency groups. The manipulation of the prime word type and phoneme type in the different tasks induced the same interference/facilitation effects for both groups. This indicated that not only did both groups experience cross-language activation, but also their responses to the various linguistic input was similar. However, the highly proficient bilinguals showed faster lexical retrieval compared to the less proficient bilinguals over all experiments. The slower processing recorded in the performance of the less proficient bilinguals has typically been attributed to weak connections between the L2 lexical forms and their concepts (Kroll et al. 2010; Van Hell and Tanner 2012). The findings of this study do not support the view that lexical selection is a dynamic process that may function in a language specific or non-specific way depending on proficiency level (e.g., Kheder and Kaan 2019). The proposed model describes the manner of lexical access and the selection processes for both proficiency groups. The importance of our findings is that they add to the body of knowledge with regard to the role of proficiency level in language production, as proficiency level did not modulate cross language activation or the manner of language selection.

9.3 The Pedagogical implications of the findings

The main findings of this study are that there is cross language activation across all linguistic levels and that the language selection process can be influenced by any available linguistic input which might increase the activation level of alternative items, which we now reflect on in reference to the implications of this finding for L2 teaching. In second language classrooms (L2), learners are often exposed to words in semantic clusters (i.e., semantically related words introduced simultaneously such as names of animals, colours, etc.), and there is an ongoing debate as to whether this is beneficial to L2 learners or if it impedes the learning process. Opponents of this semantic clustering approach argue that it hinders vocabulary learning, as it causes interference (for a review see Baddeley 1997). This is because when Arabic-English learners, for example, are introduced simultaneously to the L2 names of animals e.g., *dog* and *cat*, the learner might become confused about which word refers to which animal. Considering the findings of cross-language activation, we hypothesize that not only within-language interference will be found in a given scenario, but also cross-language interference could take place if L1 nodes receive extra activation from available linguistic input. Examples of the scenarios where alternatives could be highly active occur when teachers use L1 names to introduce the meaning of new items of vocabulary or provide a semantically related list of words in L1 and ask learners to translate the words into L2. Therefore, it is not only within language interreference that arises as a possible outcome when using semantic clusters, but there is also a chance of cross-language interference, which will then impede the learning process. Thus, the findings of this study suggest semantic clustering is not the optimum way to teach vocabulary items.

However, cross-language activation might be expected to influence the learning process positively when using cognates. Former research on L2 vocabulary acquisition argued in favour of teaching cognate words to learners to facilitate both word recognition and comprehension (e.g., Pérez et al. 2010), as well as improve learning strategies (e.g., Mugford 2008), boost self-confidence, and enhance vocabulary competence (Wallace 2007). Here we argue that the production of cognate vocabulary is faster and more accurate than non-cognate vocabulary, as per the findings of our study (cf. experiments one [chapter four]). Thus, introducing cognates into L2 classes could improve the production of L2 vocabulary and ultimately bolster learners' self-confidence. For example, L2 learners in a simple conversation task, or in an L2 naming task, would benefit from including additional cognate names/words as prompt words. This would also facilitate their lexical access and keep them engaged in the task, and motivate the less proficient learners to participate in a task, since their L1 knowledge (accessed through L1 cognates) would help them to complete it successfully.

In view of our findings, it is apparent that there is a need to carefully assess the feasibility of using L1 in L2 classrooms. Arguably, using L1 in L2 classes may facilitate access to concepts, at the same time as strengthening the lexical links between L2 and L1, because access to concept can only be achieved through L1 lexical nodes (RHM Kroll and Stewart 1994). Thus, teachers are encouraged to use L1 in L2 classes to explain new or difficult concepts. They consider it a beneficial approach to achieve the objectives of the lessons. However, during production, the use of L1 should be carefully considered, as instead of facilitating the production process, an interference would manifest. As discussed previously, our findings suggest that the process of language selection is vulnerable to any available linguistic input especially if it has a high level of activation. Depending on the type of linguistic cue available, the production process is differently affected. For example, if the learning task includes semantically related clusters, then the production process might be impeded, but if the task includes cognates, the production process could be facilitated.

9.4 Evaluation of the masked priming method

This study employed the masked priming method in relation to picture naming (experiments one, two and three) and examined the contradictory findings in the literature, as discussed in section 2.8. The procedure successfully revealed the underlying processing of word production and provided informative data concerning lexical access, the manner of lexical/phonological selection, and the flow of activation. We were able to test for the commonly reported cognate facilitation effect, semantic interference effect, and identity effect, which form the main body of evidence for current bilingual lexical access theories. Since this is the first attempt to replace the picture-word interference task with the masked priming paradigm to test current theories, further investigations are recommended, especially with same /different script bilinguals, as discussed in the next section.

9.5 Limitation and Recommendation

The present study has made important contributions to current understanding, and promotes further research on bilingual word production. However, it does have some limitations. First, it is always preferable to unify data collection methods, especially if the research is addressing the same questions; however, the outbreak of the Covid-19 virus forced us to shift from lab-based data collection to online data collection. Researchers have demonstrated that data provided by web-based tools are as reliable as those collected in lab settings (e.g., Gosling et al. 2004). However, there are reported disadvantages to be considered. Firstly, participants' environments in web-based experiments are not as tightly controlled as in traditional lab-experiments. For example, participants may not pay as much attention to the task in hand as they would in a controlled environment as there are likely to be distractors (e.g., family members, tv, etc.) around them that might influence responses. Therefore, researchers ask participants to stay in a quiet room and to pay attention to the task in hand but there is no way to check for compliance with these instructions. Secondly, there has been some reported delay of the stimulus presentation times and the recorded reaction times. For example, researchers (Reimers and Stewart 2015; de Leeuw and Motz 2016; Hilbig 2016; Semmelmann and Weigelt 2017) compared recorded reaction times in a traditional lab-set-up with a web-based collection and found delayed reaction times for the web-based method which had an impact on the size of the observed differences between conditions in the between-participants designs but not in within-participants designs. However, Reimers & Stewart (2015) suggested that the effect caused by delayed reaction times in between-participants designs can be compensated by increasing the number of participants in each group which is what we did in this study. For future work, we recommend using the online method as a second option, in case lab-based data collection is not possible. In lab-based data collection, the researcher would have better control over the external variables that influence participants when performing tasks (e.g., weak internet connections, etc) which is a component missing from the online method.

Second, the Arabic-English bilinguals in this study were all females, as no male participants were accessible at the time of the study. We cannot be certain of whether the same findings would have been obtained from a mixed population. Third, the present study investigated cross language activation, and manner of lexical selection in different script bilinguals using the masked priming paradigm. It is critical to use the same paradigm when testing same script bilinguals, or to compare both same and different script bilinguals using the same paradigm. The masked priming paradigm revealed similar results to those obtained from the picture word interference tasks, yet it is important to determine whether the same findings would be obtained with a different group of bilinguals regardless of difference in scripts. Furthermore, in future research, it is will be important to investigate the time course of lexical selection for both same and different script bilinguals. In the present research we manipulated the SOA values for (the time of prime word presentations) only in experiment two when looking for a semantic interference effect in this priming paradigm, as it was unclear at what SOA this effect would manifest. Finally, no studies have yet investigated the effect of the nature of the linguistic environment in different-script bilingual studies; namely whether living in an L2 environment with greater exposure to L2 would prevent access to L1 during production, as it is less commonly used in such a situation compared to in an L1 dominant environment.

9.6 Conclusion

This study attempted to investigate lexical access in different script bilinguals, namely Arabic-English bilinguals, and establish whether the flow of activation cascades to the phonological level or not. The results from the five consecutive experiments revealed that different script bilinguals experience non-selective lexical access that is similar to same script bilinguals, and that the flow of activation is cascaded; i.e., non-target phonological representations were active at the phonological level. The second objective was to examine the manner of lexical selection (language specific vs non-specific). The overall findings suggest that the manner of lexical/phonological selection cannot be described in terms of specificity nor non-specificity. Finally, the study analysed the effect of proficiency level on the manner of lexical access and selection and the flow of activation. The findings indicated that proficiency level did not modulate the manner of lexical access or the flow of activation, as both proficiency groups experienced non-selective access and cascaded activation of non-target phonological representations at the phonological level. Moreover, the manner of lexical selection was consistent across both groups.

Based on these findings a model is proposed in which the selection mechanism is in compliance with the preverbal message (i.e., the picture and the task instructions (to retrieve L2 name)). However, this selection mechanism is open to any available linguistic cues; thus, it would verify activated non-target lexical nodes at the lexical level. The duration of this verification process is dependent on the number of activated non-target nodes activated (i.e., the wider the pool of selection, the longer the processing time). Moreover, we concurred that the selection process at the lexical level is not finial, as the selection is subject to change at the phonological level when encountering linguistic cues at that level. The findings reported in this study could inform future studies in the area of L2 lexical access and lexical selection, as well as contribute to the general body of research into different script bilinguals.

9.7 Works Cited:

Abdel Rahman, R. and Aristei, S. 2010. Now you see it … and now again: semantic interference reflects lexical competition in speech production with and without articulation. *Psychonomic Bulletin & Review* 17(5), pp. 657-661. doi: 10.3758/PBR.17.5.657

Alario, F., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. and Segui, J. 2004. Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers* 36(1), pp. 140-155. doi: 10.3758/bf03195559

Alario, F., Segui, J. and Ferrand, L. 2000. Semantic and associative priming in picture naming. *Q J Exp Psychol A* 53(3), pp. 741-764. doi: 10.1080/713755907

AlMahmoud, M. 2013. Discrimination of Arabic contrasts by American learners. *Studies in Second Language* 3(2), pp. 261-292.

Alosh, M. 1987. *The perception and acquisition of pharyngealized fricatives by American learners of Arabic and implications for teaching Arabic phonology.* PhD Thesis, Ohio State University.

Alotaibi, Y. and Meftah, A. 2013. Review of distinctive phonetic features and the Arabic share in related modern research. *Turkish Journal of Electrical Engineering and Computer Sciences* 21(5), pp. 1426-1439.

Azuma, T. and Van Orden, C. 1997. Why safe is better than fast: the relatedness of a word's meanings affects lexical decision times. *Journal of Memory and Language* 36(4), pp. 484-504.

Baddeley, A. 1997. *Human memory: theory and practice (Revised Edition)*. East Sussex: Psychology Press.

Bajo, M., Puerta-Melguizo, M. and Macizo, P. 2003. The locus of semantic interference in picture naming. *Psicológica* 24(1), pp. 31-55. doi: 10.1037/a0033745

Blackwood, D. and Muir, W. 1990. Cognitive brain potentials and their application. *British Journal of Psychiatry* 157(S9), pp. 96-101. doi: 10.1192/S0007125000291897

Blanco-Elorrieta, E. and Caramazza, A. 2021. A common selection mechanism at each linguistic level in bilingual and monolingual language production. *Cognition* 2*13*, pp. 2-15. doi: org/10.1016/j.cognition.2021.104625

Blumenfeld, H. and Marian, V. 2007. Constraints on parallel activation in bilingual spoken language processing: examining proficiency and lexical status using eye-tracking. *Language and Cognitive Processes* 22(5), pp. 633-660. doi: 10.1080/01690960601000746

Bock, K. and Levelt, W. 1994. Language production: grammatical encoding. In: Gernsbacher, M.A. ed. *Handbook of psycholinguistics*. San Deigo, CA: Academic Press, Inc, pp. 945-984.

Boudelaa, S. and Marslen-Wilson, W. 2010. Aralex: a lexical database for Modern Standard Arabic. *Behavior Research Methods* 42(2), pp. 481-487. doi: 10.3758/BRM.42.2.481

Boukadi, M., Davies, R. and Wilson, M. 2015. Bilingual lexical selection as a dynamic process: evidence from Arabic-French bilinguals. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* 69(4), pp. 297-313. doi: 10.1037/cep0000063

Bridges, D., Pitiot, A., MacAskill, M. and Peirce, J. 2020. The timing mega-study: comparing a range of experiment generators, both lab-based and online. *PeerJ* 8, pp. e9414-e9414. doi: 10.7717/peerj.9414

Brysbaert, M., Wijnendaele, I. and Deyne, S. 2000. Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica* 104(2), pp. 215-226. doi: 10.1016/S0001-6918(00)00021-4

Bugaiska, A., Grégoire, L., Camblats, A. M., Gelin, M., Méot, A. and Bonin, P. 2019. Animacy and attentional processes: evidence from the Stroop task. *Q J Exp Psychol (Hove)* 72(4), pp. 882-889. doi: 10.1177/1747021818771514

Calabria, M., Hernández, M., Branzi, F. and Costa, A. 2012. Qualitative differences between bilingual language control and executive control: evidence from task-switching. *Frontier in Psychology* 2, p. 399. doi: 10.3389/fpsyg.2011.00399

Caramazza, A. 1997. How many levels of processing are there in lexical access? *Cognitive Neuropsychology* 14(1), pp. 177-208. doi: 10.1080/026432997381664

Chauncey, K., Holcomb, P. and Grainger, J. 2009. Primed picture naming within and across languages: an ERP investigation. *Cognitive, Affective & Behavioral Neuroscience* 9(3), pp. 286-303. doi: 10.3758/CABN.9.3.286

Chen, J., Chen, T. and Dell, G. 2002. Word-form encoding in Mandarin Chinese as assessed by the implicit priming task. *Journal of Memory and Language* 46(4), pp. 751-781. doi: 10.1006/jmla.2001.2825

Chenu, F. and Jisa, H. 2009. Reviewing some similarities and differences in L1 and L2 lexical development. *AILE (Acquisition et Interaction en Langue Etrangère)* 9(1), pp. 17-38. doi: 10.4000/aile.4506

Chéreau, C., Gaskell, M. and Dumay, N. 2007. Reading spoken words: orthographic effects in auditory priming. *Cognition* 102(3), pp. 341-360.

Cloud, N., Genesee, F. and Hamayan, E. 2000. *Dual language instruction: a handbook for enriched education*. Boston: Heinle & Heinle.

Cohen, J. 1988. *Statistical power analysis for the behavioral Sciences*. New York: NY: Routledge Academic.

Collina, S., Tabossi, P. and De Simone, F. 2013. Word production and the picture-word interference paradigm: the role of learning. *J Psycholinguist Res* 42(5), pp. 461-473. doi: 10.1007/s10936-012-9229-z

Collins, A. and Ellis, A. 1992. Phonological priming of lexical retrieval in speech production. *British Journal of Psychology* 83(3), pp. 375-388. doi: 10.1111/j.2044-8295.1992.tb02448.x

Colomé, À. 2001. Lexical activation in bilinguals' speech production: Language-specific or language-independent? *Journal of Memory and Language* 45(4), pp. 721-736. doi: 10.1006/jmla.2001.2793

Coltheart, M. 1981. The MRC psycholinguistic database. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology* 33A(4), pp. 497-505. doi: 10.1080/14640748108400805

Costa , A. 2004. Speech production in bilinguals. In: Bhatia, T.K. and Ritchie, W.C. eds. *The handbook of bilingualism*. USA: Blackwell Publishing Ltd, pp. 201-223.

Costa, A. 2005. Lexical access in bilingual production. In: Kroll, J. and de Groot, A. eds. *Handbook of bilingualism: psycholinguistic approaches*. New York, US: Oxford University Press, pp. 289-307.

Costa, A., Alario, F. and Caramazza, A. 2005. On the categorical nature of the semantic interference effect in the picture-word interference paradigm. *Psychon Bull Rev.* 12(1), pp. 125-131. doi: 10.3758/bf03196357

Costa, A. and Caramazza, A. 1999. Is lexical selection in bilingual speech production language-specific? Further evidence from Spanish–English and English–Spanish bilinguals. *Bilingualism: Language and Cognition* 2(3), pp. 231-244. doi: 10.1017/s1366728999000334

Costa, A., Caramazza, A. and Sebastian-Galles, N. 2000. The cognate facilitation effect: Implications for models of lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26(5), pp. 1283-1296. doi: 10.1037/0278-7393.26.5.1283

Costa, A., Colom, N., Gmez, O. and Sebastin-Galls, N. 2003. Another look at cross-language competition in bilingual speech production: lexical and phonological factors. *Bilingualism: Language and Cognition* 6(3), pp. 167-179. doi: 10.1017/s1366728903001111

Costa, A., Colomé, A. and Caramazza, A. 2000. Lexical access in speech production: the bilingual case. *Psicológica* 21(2), pp. 403-437.

Costa, A., La Heij, W. and Navarrete, E. 2006. The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition* 9(2), pp. 137-151. doi: 10.1017/s1366728906002495

Costa, A., Miozzo, M. and Caramazza, A. 1999. Lexical selection in bilinguals: do words in the bilingual's two lexicons compete for selection? *Journal of Memory and Language* 41(3), pp. 365-397. doi: 10.1006/jmla.1999.2651

Costa, A. and Santesteban, M. 2004. Lexical access in bilingual speech production: evidence from language switching in highly proficient bilinguals and L2 learners. *Journal of Memory and Language* 50(4), pp. 491-511. doi: 10.1016/j.jml.2004.02.002

Costa, A., Santesteban, M. and Ivanova, I. 2006. How do highly proficient bilinguals control their lexicalization process? Inhibitory and language-specific selection mechanisms are both functional. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 32(5), pp. 1057-1074. doi: 10.1037/0278-7393.32.5.1057

Cutting, J. and Ferreira, V. 1999. Semantic and phonological information flow in the production lexicon. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 25(2), pp. 318-344. doi: 10.1037/0278-7393.25.2.318

De Bot, K. 1992. A Bilingual production model: Levelt's 'speaking' model adapted. *Applied Linguistics* 13(1), pp. 1-24.

de Leeuw, J. and Motz, B. 2016. Psychophysics in a web browser? Comparing response times collected with JavaScript and Psychophysics Toolbox in a visual search task. *Behav Res Methods* 48(1), pp. 1-12. doi: 10.3758/s13428-015-0567-2

Dell, G. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review* 93(3), pp. 283-321. doi: 10.1037/0033-295X.93.3.283

Dell, G., Schwartz, F., Martin, N., Saffran, E. and Gagnon, D. 1997. Lexical access in aphasic and nonaphasic speakers. *Psychological Review* 104(4), pp. 801-838. doi: 10.1037/0033-295X.104.4.801

Dufour, S. 2008. Phonological priming in auditory word recognition: when both controlled and automatic processes are responsible for the effects. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale* 62(1), pp. 33-41. doi: 10.1037/1196-1961.62.1.33

Ferrand, L., Grainger, J. and Segui, J. 1994. A study of masked form priming in picture and word naming. *Mem Cognit.* 22(4), pp. 431-441. doi: 10.3758/bf03200868

Field, A. 2013. *Discovering statistics using IBM SPSS statistics*. Fourth ed. Los Angeles: Sage Publications.

Finkbeiner, M. and Caramazza, A. 2006. Now You See it, Now you Don't: On Turning Semantic Interference Into Facilitation in a Stroop-Like Task. *Cortex* 42(6), pp. 790-796. doi: 10.1016/s0010-9452(08)70419-2

Flege, J. 1987. The production of 'new' and 'similar' sounds in a foreign language: evidence for the effect of bilingual classification. *Journal of Phonetics* 15(1), pp. 47–65.

Fliessbach, K., Buerger, C., Trautner, P., Elger, E. and Weber, B. 2010. Differential effects of semantic processing on memory encoding. *Human brain mapping* 31(11), pp. 1653-1664. doi: 10.1002/hbm.20969

Forster, K. and Davis, C. 1984. Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 10(4), pp. 680-698. doi: 10.1037/0278-7393.10.4.680

Frauenfelder, U., Segui, J. and Dijkstra, T. O. N. 1990. Lexical effects in phonemic processing: facilitatory or inhibitory? *Journal of Experimental Psychology: Human Perception and Performance* 16(1), pp. 77-91. doi: 10.1037/0096-1523.16.1.77

Gelin, M., Bugaiska, A., Méot, A. and Bonin, P. 2017. Are animacy effects in episodic memory independent of encoding instructions? *Memory* 25(1), pp. 2-18. doi: 10.1080/09658211.2015.1117643

Glaser, W. and Düngelhoff, F. 1984. The time course of picture-word interference. *Journal of Experimental Psychology: Human Perception and Performance* 10(5), pp. 640-654. doi: 10.1037/0096-1523.10.5.640

Gollan, T. and Acenas, L. 2000. Tip-of-the-tongue incidence in Spanish-English and Tagalog-English bilinguals. *Proceedings of the Third International Symposium on Bilingualism.* Bristol, U.K.

Gollan, T. and Kroll, J. 2001. Bilingual lexical access. In: Rapp, B. ed. *The handbook of cognitive neuropsychology: what deficits reveal about the human mind*. New York, NY, US: Psychology Press, pp. 321-345.

Gosling, S. and Mason, W. 2015. Internet research in psychology. *Annu Rev Psychol* 66(1), pp. 877-902. doi: 10.1146/annurev-psych-010814-015321

Gosling, S., Vazire, S., Srivastava, S. and John, O. 2004. Should we trust web-based studies? A Comparative analysis of six preconceptions about internet questionnaires. *American Psychologist 59*(2), pp. 93–104. doi: 10.1037/0003-066X.59.2.93

Green, D. 1986. Control, activation, and resource: a framework and a model for the control of speech in bilinguals. *Brain Lang* 27(2), pp. 210-223. doi: 10.1016/0093-934X(86)90016-7

Green, D. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition* 1(2), pp. 67-81. doi: 10.1017/S1366728998000133

Grosjean, F. 2001. The Bilingual's Language Modes. In: Nicol, J.L. ed. *One Mind, Two Languages: Bilingual Language Processing.* Malden: Blackwell Publishing, pp. 1-22.

Grosjean, F. and Li, P. 2013. *The psycholinguistics of bilingualism*. Oxford, England: Wiley-Blackwell.

Han, J. and Choi, T. 2016. The influence of spelling on the production and storage of words with allophonic variants of /h/ in Korean. *Applied Psycholinguistics* 37(4), pp. 757-780. doi: 10.1017/S0142716415000235

Harley, T. 2001. *Psychology of language: from data to theory*. New York: Psychology Press Ltd.

Hermans, D. 2000. *Word production in a foreign language.* PhD Thesis, University of Nijmegen.

Hermans, D. 2004. Between-language identity effects in picture-word interference tasks: a challenge for language-nonspecific or language-specific models of lexical access? 8(2), pp. 115-125. doi: 10.1177/13670069040080020101

Hermans, D., Bongaerts, T., De Bot, K. and Schreuder, R. 1998. Producing words in a foreign language: Can speakers prevent interference from their first language? *Bilingualism: Language and Cognition* 1(3), pp. 213-229. doi: 10.1017/s1366728998000364

Hermans, D., Ormel, E., Besselaar, R. and Hell, J. 2011. Lexical activation in bilinguals' speech production is dynamic: how language ambiguous words can affect cross-language activation. *Language and Cognitive Processes* 26(10), pp. 1687-1709. doi: 10.1080/01690965.2010.530411

Hilbig, B. 2016. Reaction time effects in lab-versus web-based research: experimental evidence. *Behav Res Methods* 48(4), pp. 1718-1724. doi: 10.3758/s13428-015-0678-9

Hoshino, N. 2006. *A Psycholinguistic study of native language constraints on speaking words in a second language.* PhD Thesis, Pennsylvania State University.

Hoshino, N., Beatty-Martínez, A., Navarro-Torres, C. and Kroll, J. 2021. Do cross-language script differences enable bilinguals to function selectively when speaking in one language alone? *Frontiers in Communication* 6, doi: 10.3389/fcomm.2021.668381

Hoshino, N. and Kroll, J. 2008. Cognate effects in picture naming: does cross-language activation survive a change of script? *Cognition* 106(1), pp. 501-511. doi: 10.1016/j.cognition.2007.02.001

Huibregtse, I., Admiraal, W. and Meara, P. 2002. Scores on a yes-no vocabulary test: correction for guessing and response style. *Language Testing* 19(3), pp. 227-245. doi: 10.1191/0265532202lt229oa

Humphreys, G. W., Riddoch, M. J. and Quinlan, P. T. 1988. Cascade processes in picture identification. *Cognitive Neuropsychology* 5(1), pp. 67-104. doi: 10.1080/02643298808252927

Jackson, G., Swainson, R., Cunnington, R. and Jackson, S. 2001. ERP correlates of executive control during repeated language switching. *Bilingualism: Language and Cognition* 4(02), pp. 169-178. doi: 10.1017/s1366728901000268

Jacobs, A., Fricke, M. and Kroll, J. 2016. Cross-language activation begins during speech planning and extends into second language speech. *Lang Learn* 66(2), pp. 324-353. doi: 10.1111/lang.12148

Janssen, N. 1999. *Bilingual word production: The time course of lexical activation in a mixed language context.* University of Nijmegen.

Janssen, N., Schirm, W., Mahon, B. and Caramazza, A. 2008. Semantic interference in a delayed naming task: evidence for the response exclusion hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34(1), pp. 249-256. doi: 10.1037/0278-7393.34.1.249

Jared, D. and Kroll, J. 2001. Do bilinguals activate phonological representations in one or both of their languages when naming words? *Journal of Memory and Language* 44(1), pp. 2-31. doi: 10.1006/jmla.2000.2747

Jescheniak, J. and Schriefers, H. 1998. Discrete serial versus cascaded processing in lexical access in speech production: further evidence from the coactivation of near-synonyms. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24(5), pp. 1256-1274. doi: 10.1037/0278-7393.24.5.1256

Kheder, S. and Kaan, E. 2019. Lexical selection, cross-language interaction, and switch costs in habitually codeswitching bilinguals. *Bilingualism: Language and Cognition* 22(3), pp. 569-589. doi: 10.1017/S1366728918000500

Khwaileh, T., Mustafawi, E., Herbert, R. and Howard, D. 2018. Gulf Arabic nouns and verbs: a standardized set of 319 object pictures and 141 action pictures, with predictors of naming latencies. *Behav Res Methods* 50(6), pp. 2408-2425. doi: 10.3758/s13428-018-1019-6.

Kochari, A. 2019. Conducting web-based experiments for numerical cognition research. *J Cogn* 2(1), p. 39. doi: 10.5334/joc.85

Kroll, J., Bobb, S., Misra, M. and Guo, T. 2008. Language selection in bilingual speech: evidence for inhibitory processes. *Acta Psychol (Amst)* 128(3), pp. 416-430. doi: 10.1016/j.actpsy.2008.02.001

Kroll, J., Bobb, S. and Wodniecka, Z. 2006. Language selectivity is the exception, not the rule: arguments against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition* 9(2), pp. 119-135. doi: 10.1017/S1366728906002483

Kroll, J., Dijkstra, A., Janssen, N. and Schriefers, H. 2000. Selecting the language in which to speak: experiments on lexical access in bilingual production. *41st Annual Meeting of the Psychonomic Society.* New Orleans, LA.

Kroll, J., Dussias, P., Bogulski, C. and Kroff, J. 2012. Juggling two languages in one mind: what bilinguals tell us about language processing and its consequences for cognition. In: Ross, B. ed. *The psychology of learning and motivation, Vol. 56.* San Diego, CA, US: Elsevier Academic Press, pp. 229-262.

Kroll, J. and Stewart, E. 1994. Category interference in translation and picture naming - evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33(2), pp. 149-174. doi: DOI 10.1006/jmla.1994.1008

Kroll, J., Sumutka, B. and Schwartz, A. 2005. A cognitive view of the bilingual lexicon: reading and speaking words in two languages. *International Journal of Bilingualism* 9, pp. 27- 48.

Kroll, J. and Sunderman, G. 2003. Cognitive processes in second language acquisition: the development of lexical and conceptual representations. In: Doughty, C.J. and Long, M.H. eds. *Handbook of second language acquisition*. Cambridge, MA: Blackwell Publishers, pp. 104-129.

Kroll, J., van Hell, J., Tokowicz, N. and Green, D. 2010. The revised hierarchical model: a critical review and assessment. *Biling (Camb Engl)* 13(3), pp. 373-381. doi: 10.1017/S136672891000009X

La Heij, W. 1988. Components of Stroop-like interference in picture naming. *Mem Cognit.* 16(5), pp. 400-410. doi: 10.3758/bf03214220. PMID: 3173088.

La Heij, W. 2005. Selection processes in monolingual and bilingual lexical access. In: Kroll, J. and de Groot, A. eds. *Handbook of bilingualism: psycholinguistic approaches*. New York, US: Oxford University Press, pp. 289-307.

Levelt, W. 1989. *Speaking: from intention to articulation*. Cambridge, MA, US: The MIT Press.

Levelt, W., Roelofs, A. and Meyer, A. 1999. A theory of lexical access in speech production. *Behav Brain Sci* 22(1), pp. 1-38; discussion 38-75.

Levelt, W., Schriefers, H., Vorberg, D., Meyer, A., Pechmann, T. and Havinga, J. 1991. The time course of lexical access in speech production: a study of picture naming. *Psychological Review* 98(1), pp. 122-142. doi: 10.1037/0033-295X.98.1.122

Lupker, S. J. 1982. The role of phonetic and orthographic similarity in picture-word interference. *Canadian Journal of Psychology* 36, p. 349_376.

Mahon, B., Costa, A., Peterson, R., Vargas, K. and Caramazza, A. 2007. Lexical selection is not by competition: a reinterpretation of semantic interference and facilitation effects in the picture-word interference paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33(3), pp. 503-535. doi: 10.1037/0278-7393.33.3.503

McLeod, S. 2019. *What does effect size tell you?* Available at: https://www.simplypsychology.org/effect-size.html [Accessed: 17 June 2021].

McNamara, J. and Kushnir, L. 1972. Linguistic independence of bilinguals: the input switch. *Journal of Verbal Learning and Verbal Behavior* 10(5), pp. 480-487.

Meuter, I. and Tan, C. 2003. A comparative study of the role of language specificity and language mode. *Proceedings of the Fourth International Symposium on Bilingualism.* Tempe, AZ.

Meuter, R. and Allport, A. 1999. Bilingual language switching in naming: asymmetrical costs of language selection. *Journal of Memory and Language* 40(1), pp. 25-40. doi: 10.1006/jmla.1998.2602

Miller, K. and Swick, D. 2003. Orthography influences the perception of speech in alexic patients. *Journal of Cognitive Neuroscience* 15(7), pp. 981-990. doi: 10.1162/089892903770007371

Miller, N. 2001. *Language and lexical selection in Dutch/English bilinguals.* Pennsylvania State University.

Miller, N. and Kroll, J. 2002. Stroop effects in bilingual translation. *Memory & Cognition* 30(4), pp. 614-628. doi: 10.3758/BF03194963

Miozzo, M. and Caramazza, A. 2003. When more is less: A counterintuitive effect of distractor frequency in the picture-word interference paradigm. *Journal of Experimental Psychology: General* 132(2), pp. 228-252. doi: 10.1037/0096-3445.132.2.228

Mochida, A. and Harrington, M. 2006. The Yes/No test as a measure of receptive vocabulary knowledge. *Language Testing* 23(1), pp. 73-98. doi: 10.1191/0265532206lt321oa

Moon, J. and Jiang, N. 2012. Non-selective lexical access in different-script bilinguals. *Bilingualism-Language and Cognition* 15(1), pp. 173-180. doi: 10.1017/S1366728911000022

Mosca, M. and de Bot, K. 2017. Bilingual language switching: production vs. recognition. *Frontiers in Psychology* 8, pp. 1664-1078. doi: 10.3389/fpsyg.2017.00934

Mugford, G. 2008. Keeping a critical eye on "lexical friends": cognates as critical pedagogy in pre-service teacher education. *Profile Issues in Teachers' Professional Development* 9(Jan/June), pp. 129-142.

New, J., Cosmides, L. and Tooby, J. eds. 2007. *Category-specific attention for animals reflects ancestral priorities, not expertise*. *Proceedings of the National Academy of Sciences.*

Pattamadilok, C., Morais, J., Ventura, P. and Kolinsky, R. 2007. The locus of the orthographic consistency effect in auditory word recognition: further evidence from French. *Language and Cognitive Processes* 22(5), pp. 1-27.

Pecher, D. and Raaijmakers, J. 2004. Priming for new associations in animacy decision: evidence for context dependency. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology* 57A(7), pp. 1211-1231. doi: 10.1080/02724980343000738

Peirce, J. et al. 2019. PsychoPy2: experiments in behavior made easy. *Behavior Research Methods* 51(1), pp. 195-203. doi: 10.3758/s13428-018-01193-y

Pérez, A., Peña, E. and Bedore, L. 2010. Cognates facilitate word recognition in young Spanish-English bilinguals' test performance. *Early Childhood Services* 4(1), pp. 55-67.

Peterson, R. and Savoy, P. 1998. Lexical selection and phonological encoding during language production: evidence for cascaded processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24(3), pp. 539-698.

Philipp, A., Gade, M. and Koch, I. 2007. Inhibitory processes in language switching: evidence from switching language-defined response sets. *European Journal of Cognitive Psychology* 19(3), pp. 395-416. doi: 10.1080/09541440600758812

Poarch, G. and Van Hell, J. 2012. Cross-language activation in children's speech production: evidence from second language learners, bilinguals, and trilinguals. *Exp Child Psychol* 111(3), pp. 419-438. doi: 10.1016/j.jecp.2011.09.008.

Potter, M., So, K., Voneckardt, B. and Feldman, L. 1984. Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior* 23(1), pp. 23-38. doi: 10.1016/S0022-5371(84)90489-4

Poulisse, N. 1997. Language production in bilinguals. In: de Groot, A. and Kroll, J. eds. *Tutorials in bilingualism: psycholinguistic perspectives*. Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers, pp. 201-224.

Poulisse, N. and Bongaerts, T. 1994. First language use in second language production. *Applied Linguistics* 15(1), pp. 36-57. doi: 10.1093/applin/15.1.36

Pratt, J., Radulescu, P., Guo, R. and Abrams, R. 2010. It’s alive! Animate motion captures visual attention. *Psychological Science* 21(11), pp. 1724-1730. doi: 10.1177/0956797610387440

Prebianca, G. V. V. 2014. Exploring the relationship between lexical access and proficiency level in L2 speech production. *Trabalhos em Linguística Aplicada* 53(2), pp. 381-402. doi: 10.1590/s0103-18132014000200007

Qu, Q. and Damian, M. 2016. Orthographic effects in spoken word recognition: evidence from Chinese. *Psychon Bull Rev* 24(3), pp. 901-906. doi: 10.3758/s13423-016-1164-9. PMID: 27604494

Räling, R., Hanne, S., Schröder, A., Keßler, C. and Wartenburger, I. 2017. Judging the animacy of words: the influence of typicality and age of acquisition in a semantic decision task. *Quarterly Journal of Experimental Psychology* 70(10), pp. 2094-2104. doi: 10.1080/17470218.2016.1223704

Rammuny, R. 1976. Statistical study of errors made by American students in written Arabic. *Journal of American Association of Teachers of Arabic* 10(1), pp. 11-29.

Rayner, K. and Springer, C. 1986. Graphemic and semantic similarity effects in the picture-word interference task. *British Journal of Psychology* 77(2), pp. 207-222. doi: 10.1111/j.2044-8295.1986.tb01995.x

Reimers, S. and Stewart, N. 2015. Presentation and response timing accuracy in Adobe Flash and HTML5/JavaScript Web experiments. *Behavior Research Methods* 47(2), pp. 309-327. doi: 10.3758/s13428-014-0471-1

Roelofs, A. 1992. A spreading-activation theory of lemma retrieval in speaking. *Cognition* 42(1-3), pp. 107-142. doi: 10.1016/0010-0277(92)90041-F

Roelofs, A. 1998. Lemma selection without inhibition of languages in bilingual speakers. *Bilingualism: Language and Cognition* 1(02), pp. 94-95. doi: 10.1017/s1366728998000194

Roelofs, A. 2006. The influence of spelling on phonological encoding in word reading, object naming, and word generation. *Psychon Bull Rev* 13(1), pp. 33-37. doi: 10.3758/bf03193809

Roelofs, A., Meyer, A. and Levelt, W. 1998. A case for the lemma/lexeme distinction in models of speaking: comment on Caramazza and Miozzo. *Cognition* 69(2), pp. 219-230. doi: 10.1016/s0010-0277(98)00056-0. PMID: 9894405.

Schriefers, H., Meyer, A. and Levelt, W. 1990. Exploring the time course of lexical access in language production: picture-word interference studies. *Journal of Memory and Language* 29(1), pp. 86-102. doi: 10.1016/0749-596x(90)90011-N

Schwartz, A. and Kroll, J. 2006. Bilingual lexical activation in sentence context. *Journal of Memory and Language* 55(2), pp. 197-212. doi: 10.1016/j.jml.2006.03.004

Schwieter, J. and Sunderman, G. 2008. Language switching in bilingual speech production: in search of the language-specific selection mechanism. *The Mental Lexicon* 3(2), pp. 214-238. doi: 10.1075/ml.3.2.06sch

Segalowitz, N. and Hulstijn, J. 2005. Automaticity in bilingualism and second language learning. In: Kroll, J. and de Groot, A. eds. *Handbook of bilingualism: psycholinguistics approaches*. Oxford: Oxford University Press, pp. 179-201.

Semmelmann, K. and Weigelt, S. 2017. Online psychophysics: reaction time effects in cognitive experiments. *Behav Res Methods* 49(4), pp. 1241-1260. doi: 10.3758/s13428-016-0783-4

Shehata, A. 2015. Problematic Arabic consonants for native English speakers: learners' perspectives. *International Journal of Educational Investigations* 2(9), pp. 24-74.

Simonet, M. 2016. The phonetics and phonology of bilingualism. *Oxford handbooks online: scholarly research reviews*. Oxford, UK: Oxford University Press.

Starreveld, P. 2000. On the interpretation of phonological context effects in word production. *Journal of Memory and Language* 42(4), pp. 497-525. doi: 10.1006/jmla.1999.2693

Starreveld, P. and La Heij, W. 1995. Semantic interference, orthographic facilitation, and their interaction in naming tasks. *Journal of Experimental Psychology-Learning Memory and Cognition* 21(3), pp. 686-698. doi: 10.1037/0278-7393.21.3.686

Starreveld, P. and la Heij, W. 1996. Time-course analysis of semantic and orthographic context effects in picture naming. *Journal of Experimental Psychology-Learning Memory and Cognition* 22(4), pp. 896-918. doi: 10.1037/0278-7393.22.4.896

Stewart, N., Chandler, J. and Paolacci, G. 2017. Crowdsourcing samples in cognitive science. *Trends Cogn Sci* 21(10), pp. 736-748. doi: 10.1016/j.tics.2017.06.007

Stroop, J. R. 1935. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology* 18(6), pp. 643-662. doi: Doi 10.1037//0096-3445.121.1.15

Szekely, A. et al. 2003. Timed picture naming: extended norms and validation against previous studies. *Behavior Research Methods, Instruments & Computers* 35(4), pp. 621-633. doi: 10.3758/BF03195542

Szekely, A. et al. 2004. A new on-line resource for psycholinguistic studies. *Journal of Memory and Language* 51(2), pp. 247-250. doi: 10.1016/j.jml.2004.03.002

Taft, M., Castles, A., Davis, C., Lazendic, G. and Nguyen-Hoan, M. 2008. Automatic activation of orthography in spoken word recognition: pseudohomograph priming. *Journal of Memory and Language* 58(2), pp. 366-379.

Uyeda, K. and Mandler, G. 1980. Prototypicality norms for 28 semantic categories. *Behavior Research Methods & Instrumentation* 12, pp. 587-595. doi: 10.3758/BF03201848

Van Hell, J. and Tanner, D. 2012. Second language proficiency and cross-language lexical activation. *Language Learning* 62(2), pp. 148-171. doi: 10.1111/j.1467-9922.2012.00710.x

Ventura, P., Morais, J., Pattamadilok, C. and Kolinsky, R. 2004. The locus of the orthographic consistency effect in auditory word recognition. *Language and Cognitive Processes* 19(1), pp. 57-95. doi: 10.1080/01690960344000134

Verhoef, K. M., Roelofs, A. and Chwilla, D. J. 2010. Electrophysiological evidence for endogenous control of attention in switching between languages in overt picture naming. *J Cogn Neurosci* 22(8), pp. 1832-1843. doi: 10.1162/jocn.2009.21291

Wallace, C. 2007. Vocabulary: the key to teaching English language learners to read. *Education Digest: Essential Readings Condensed for Quick Review* 73(4), pp. 36-39.

Weekes, B., Davies, R. and Chen, M. 2002. *Picture-word interference effects on naming in Chinese*. Hong Kong: Hong Kong University Press.

Wheeldon, L. and Levelt, W. 1995. Monitoring the time-course of phonological encoding. *Journal of Memory and Language* 34(3), pp. 311-334. doi: DOI 10.1006/jmla.1995.1014

White, A. 2015. The cognate boost: a study of picture naming across proficiency levels with L2 learners of Russian. *Journal of Slavic Linguistics* 23(2), pp. 285-311.

Woods, A., Velasco, C., Levitan, C., Wan, X. and Spence, C. 2015. Conducting perception research over the internet: a tutorial review. *PeerJ* 3, p. e1058. doi: 10.7717/peerj.1058

Zhao, H., La Heij, W. and Schiller, N. 2012. Orthographic and phonological facilitation in speech production: new evidence from picture naming in Chinese. *Acta Psychologica* 139(2), pp. 272-280. doi: 10.1016/j.actpsy.2011.12.001

Zou, L., Desroches, A., Liu, Y., Xia, Z. and Shu, H. 2012. Orthographic facilitation in Chinese spoken word recognition: an ERP study. *Brain and Language* 123(3), pp. 164-173. doi: 10.1016/j.bandl.2012.09.006

9.8 Appendixes

APPENDIX A

CONSENT FORM

Title of research project:

Lexical access by different script adult bilinguals; Evidence from masked primed picture naming and phoneme monitoring tasks.

SREC reference:

Name of Chief/Principal Investigator:

	Please initial box
I confirm that I have read the information sheet dated [30/09/2020] version [August 2019] for the above research project.	
I confirm that I have understood the information sheet dated [30/09/2020] version [August 2019] for the above research project and that I have had the opportunity to ask questions and that these have been answered satisfactorily.	
I understand that my participation is voluntary and I am free to withdraw any time before the end of the 7-day grace period for withdrawal without giving a reason and without any adverse consequences (e.g. to my grades or my relationship with the researcher as a student/ colleague). I understand that there will be no negative impact on me as a result of my participation or non-participation in this project (e.g. my grades or professional relationship with the researcher). I understand that my decision will have no negative impact on the researcher or this project.	
I understand that withdrawal is possible at any point prior to the data analysis phase which starts 7 days after your completion of all the required tasks. I understand that during the analysis phase, the researcher will discard participant's number from the scripts that will be used in the analysis, and the researcher will not be able to withdraw your data since he/she will not be able to identify it.	
I understand that it is my responsibility to remember my participant's number (in lab based-tasks or web-based tasks) and to provide the researcher with this number in order to identify my data in case I decided to withdraw before the data analysis phase. I understand that without this number the researcher will not be able to identify my data and withdraw it upon my request.	
I understand that data collected during the research project may be looked at by individuals from Cardiff University or from regulatory authorities, where it is relevant to my taking part in the research project. I give permission for these individuals to have access to my data.	
I consent to the processing of my personal information [age, gender and nationality], my educational and linguistic background, vocal/button-press responses and reaction times of my responses and test scores for the purposes explained to me. I understand that such information will be held in accordance with all applicable data protection legislation and in strict confidence, unless disclosure is required by law or professional obligation.	

I understand who will have access to personal information provided, how the data will be stored and what will happen to the data at the end of the research project.	
I understand that the anonymized data (my personal information: age, gender, nationality educational and linguistic background, and the anonymized data extracted from my voice recordings/button-press responses and reaction times) will be used in this research project, future journal publications and conference presentations in psycholinguistic studies. I understand that vocal responses will be held confidentially and accessible only by the researcher and will be kept for a period of 3 months after commencing the tasks and then will be permanently deleted. I understand that it will be stored under the participant’s ID number in the researcher’s space on Microsoft OneDrive (a secured-password protected drive provided by Cardiff University). I understand that a score sheet will be used to represent the data collected from vocal recordings and these records will be anonymously processed in the data analysis phase, later stages of this research project, future journal publications and presentations.	
I understand that all information will be stored in such a way that it cannot be traced back to me. I understand the research team will anonymise the data (my personal information: age, gender, nationality, educational and linguistic background, button-press responses and reaction times) they have collected from, or about, me in connection with this research project from the point I perform the tasks with the exception of my signed consent form. I understand that my vocal responses will be stored under my participant’s ID number and will be kept for 3 months (data analysis period) and then permanently deleted. I understand that data extracted from the vocal recordings will be anonymously processed during the data analysis phase which starts 7 days after completion of the required tasks. I understand that signed consent forms cannot be anonymised. Anonymised information and the signed consent form will be kept for a minimum of 5 years after which will be permanently deleted (i.e. paper records will be shredded and digital records will be permanently deleted).	
I consent to being audio recorded for the purposes of the research project and I understand how it will be used in the research.	
I understand how the findings and results of the research project will be written up and published.	
I agree to take part in this research project.	

Name of participant (print) Date Signature

THANK YOU FOR PARTICIPATING IN OUR RESEARCH

YOU WILL BE GIVEN A COPY OF THIS CONSENT FORM TO KEEP

Appendix B

PARTICIPANT INFORMATION SHEET

Lexical access by different script adult bilinguals; Evidence from masked primed picture naming and phoneme monitoring tasks.

You are being invited to take part in a research project. Before you decide whether or not to take part, it is important for you to understand why the research is being undertaken and what it will involve. Please take time to read the following information carefully and discuss it with others, if you wish.

Thank you for reading this.

1. What is the purpose of this research project?

This research project is part of my PhD thesis. The objectives of this study are: (i) to investigate whether Arabic-English bilinguals experience co-activation of both languages during single word production, (ii) to investigate the manner and locus of language selection during production, and (iii) to examine whether proficiency level affect cross-language activation. The experiments will involve naming pictures on screen, deciding whether a picture is of a living thing or not, and whether a specific phoneme is part of the name of a picture. Your response will be recorded for later analysis.

2. Why have I been invited to take part?

You have been invited because you are an adult Saudi who can speak Arabic and English.

3. Do I have to take part?

No, your participation in this research project is entirely voluntary and it is up to you to decide whether or not to take part. If you decide to take part, we will discuss the research project with you and ask you to sign a consent form. If you decide not to take part, you do not have to explain your reasons and it will not affect your legal rights.

You are free to withdraw your consent to participate in the research project at any time prior to the data analysis phase which starts 7 days after your completion of all the required tasks, without giving a reason, even after signing the consent form. you are not negatively affected in any form as a result of your participation or non-participation in this project (e.g. your grades or professional relationship with the researcher). Please be assured that your decision will have no negative impact on the researcher or this project

4. What will taking part involve?

- Completing a language history questionnaire and it will require 15 minutes of your time.
- Completing a lexical decision task in which your reaction times and vocal/button-press responses will be recorded for research purpose. You will be asked to response YES if the word presented on the screen is an English word and NO if it is not a word. Your scores will be used as criteria to help placing your task results in the appropriate proficiency group. It will require 5 minutes of your time.

- Completing a picture naming and animacy decision task in which your vocal/button-press response and reaction times will be recorded for research purpose. It will take around 5 minutes of your time.
- Or completing a phoneme monitoring task in which your vocal/button-press responses and reaction times will be recorded for research purpose. It will take around 5 minutes of your time.
- Completing a brief training session which will take 5 minutes. You will be presented with the phonemes and their sounds to familiarize yourself with them before doing a simple task.

5. Will I be paid for taking part?

No. You should understand that any data you give will be as a gift and you will not benefit financially in the future should this research project lead to the development of a new method.

6. What are the possible benefits of taking part?

There will be no direct advantages or benefits to you from taking part, but your contribution will help us understand the mechanism behind the lexical selection process during the production of single word.

7. What are the possible risks of taking part?

There are no anticipated risks to taking part in your study, however, you will be informed before starting any task that you can stop and withdraw at any point prior to the data analysis phase which starts 7 days after your completion of all the required tasks.

8. Will my taking part in this research project be kept confidential?

All information collected from (or about) you during the research project will be kept confidential and any personal information you provide will be managed in accordance with data protection legislation. Please see 'What will happen to my Personal Data?' (below) for further information.

9. What will happen to my Personal Data?

Cardiff University is the Data Controller and is committed to respecting and protecting your personal data in accordance with your expectations and Data Protection legislation. Further information about Data Protection, including:

- your rights
- the legal basis under which Cardiff University processes your personal data for research
- Cardiff University's Data Protection Policy
- how to contact the Cardiff University Data Protection Officer
- how to contact the Information Commissioner's Office

may be found at https://www.cardiff.ac.uk/public-information/policies-and-procedures/data-protection

In this study, you will be required to fill in a language history questionnaire in which you provide us with your personal information (age, gender, and nationality) and a general educational and linguistic background to identify your proficiency level in English for this research purpose only. Your vocal/button-press responses and reaction times in the tasks will be recorded. You will be given a participant number by the researcher, if you perform lab-based tasks, to keep your identity anonymized. Whereas, if you performed online tasks, you have to create your own ID number. Please use this number when completing the questionnaire, lexical decision task, and the other experiments. All information will be stored in such a way that it cannot be traced back to you. Please note that signed consent forms cannot be anonymised. The research team will anonymise all the data (age, gender, nationality, educational and linguistic background, button-press responses and reaction times) it has collected from, or about, you in connection with this research project, from the point you do the tasks with the exception of your signed consent form. Vocal responses will be held confidentially and accessible only by the researcher and will be kept for a period of 3 months after commencing the tasks and then will be permanently deleted. It will be stored under the participant's ID number in the researcher's space on Microsoft OneDrive (a secured-password protected drive provided by Cardiff University). A scoring sheet will be used to represent the data collected from vocal recordings and these records will be anonymously processed in the data analysis phase, later stages of this research project, future journal publications and presentations.

Anonymised information and the signed consent form will be kept for a minimum of 5 years after which will be permanently deleted (i.e. paper records will be shredded and digital records will be permanently deleted).

Upon your request, the researcher can withdraw your data from the project (during the 7-day grace period for withdrawal) you just need to provide the researcher with your participant's number. If you performed the tasks online, participant's number is the ID number you created and entered at the beginning of each task. If you performed the tasks in lab, it is the ID number that the researcher will provide you with and it refers to your order of participation. The researcher will not keep any record that connect this number to you. So, you must remember this number to provide the researcher with in case you decide to withdraw form this project. Withdrawing is possible at any point prior to the data analysis phase which starts 7 days after your completion of all the required tasks. At that point, the researcher will discard participant's number from all scripts and the researcher will not be able to withdraw your data since he/she will not be able to identify it.

10. What happens to the data at the end of the research project?

Anonymized data will be used in future journal publications and conference presentations in psycholinguistic studies. Anonymised information and the signed consent form will be kept for a minimum of 5 years after which will be permanently deleted (i.e., paper records will be shredded and digital records will be permanently deleted).

11. What will happen to the results of the research project?

The results will be used for a University assessment and are likely to be published in academic journals and presented at conferences. A copy of the published results can be requested through contacting the researcher via email. Please note that participants will not be identified in any report, publication or presentation.

12. What if there is a problem?

If you wish to complain, or have grounds for concerns about any aspect of the manner in which you have been approached or treated during the course of this research, please contact the Supervisor, Dr. Michelle Aldridge-Waddon (AldridgeM@cardiff.ac.uk).

If your complaint is not managed to your satisfaction, please contact the ENCAP Research Ethics Officer, Dr Sara Pons-Sanz (pons-sanzs@cardiff.ac.uk).

If you are harmed by taking part in this research project, there are no special compensation arrangements. If you are harmed due to someone's negligence, you may have grounds for legal action, but you may have to pay for it.

13. Who is organising and funding this research project?

The research is organised by Manal Alharbi, a research student, and Dr. Michelle Aldridge-Waddon, the lead supervisor, affiliated to the School of English, Communication and Philosophy in Cardiff University. The research is currently funded by Jeddah University.

14. Who has reviewed this research project?

This research project has been reviewed and given a favourable opinion by the ENCAP's Research Ethics Committee through its proportionate review process.

15. Further information and contact details

Should you have any questions relating to this research project, you may contact us during normal working hours:

Manal Alharbi
alharbim1@cardiff.ac.uk

Dr. Michelle Aldridge-Waddon
AldridgeM@cardiff.ac.uk

Thank you for considering participating in this research project. If you decide to participate, you will be given a copy of the Participant Information Sheet and a signed consent form to keep for your records.

Appendix C

L2 Language History Questionnaire (Version 2.0)

Please complete this form in English or Arabic.
Experiment ID no:___________________________ Today's Date _________________

Please answer the following questions to the best of your knowledge.

PART A

1. Do you speak a second language?

YES, my second language is ____________________.

NO (If you answered NO, you need not to continue this form).

2. Do you speak more than two languages?

YES (If you answered YES, you need not to continue this form).
No

3. Age (in years): ----------------------

4. Sex (circle one): Male / Female / Prefer not to say

5 Education (degree obtained or school level attended):

6. (a). Country of origin:

(b). Country of Residence:

7. If 6(a) and 6(b) are the same, how long have you lived in a foreign country where your second language is spoken? If 6(a) and 6(b) are different, how long have you been in the country of your current residence? (in years)

---

8. What is your native language? (If you grew up with more than one language, please specify)

---

9. Please specify the age at which you started to learn your second language in the following situations (write age next to any situation that applies).

At home: __________

In school: __________

After arriving in the second language speaking country _________

10. How did you learn your second language? (check all that apply)

(Mainly Occasionally) through formal classroom instruction.

(Mainly Occasionally) through interacting with people.

A mixture of both, but (More classroom More interaction Equally both).

Other (**specify:** ____).

11. Rate your ability on the following aspects in your second language. Please rate according to the following scale (write down the number in the table):

Poor Native-like

1 _______2_________3_________4__________5

Language	**Reading proficiency**	**Writing proficiency**	**Speaking fluency**	**Listening ability**

12. Provide the age at which you were first exposed to the second language in terms of speaking, reading, and writing, and the number of years you have spent on learning your second language.

Language	**Age first exposed to the language**			**Number of years learning**
	Speaking	**Reading**	**Writing**	

PART B

13. Estimate, in terms of percentages, how often you use your native language and second language per day (in all daily activities combined, circle one that applied):

Native language: <25% 25% 50% 75% 100%

Second language: <25% 25% 50% 75% 100%

(specify the languages: ____________________)

14. Estimate, in terms of hours per day, how often you are engaged in the following activities with your native and second languages.

Activities	**First Language**	**Second Language**
Listen to Radio:	_______(hrs)	_______(hrs)
Watching TV:	_______(hrs)	_______(hrs)
Reading for fun:	_______(hrs)	_______(hrs)
Reading for work:	_______(hrs)	_______(hrs)
Reading on the Internet:	_______(hrs)	_______(hrs)
Writing emails to friends:	_______(hrs)	_______(hrs)
Writing articles/papers:	_______(hrs)	_______(hrs)

15. Estimate, in terms of hours per day, how often you speak (or used to speak) your native and second languages with the following people.

	Native Language	**Second Language**	**Hours**
Father :	________________	______________	___________ (hrs)
Mother:	________________	______________	___________ (hrs)
Grandfather (s):	________________	______________	___________ (hrs)
Grandmother(s):	________________	______________	___________ (hrs)
Brother(s)/Sister(s):	________________	______________	___________ (hrs)
Other family members	________________	______________	___________ (hrs)

16. Estimate, in terms of hours per day, how often you now speak your native and second languages with the following people.

	Native Language	**Second Language**	**Hours**
Spouse/partner:	________________	______________	___________ (hrs)
Friends:	________________	______________	___________ (hrs)
Classmates:	________________	______________	___________ (hrs)
Co-workers:	________________	______________	___________ (hrs)
Other family members	________________	______________	___________ (hrs)

17. Write down the name of the language in which you received instruction in school, for each schooling level:

Primary/Elementary School: __________

Secondary/Middle School: __________

High School: __________

College/University: __________

18. When you are speaking, do you ever mix words or sentences from the two languages you know? (If no, skip to question 19).

---

19. List the languages that you mix and rate the frequency of mixing in normal conversation with the following people according to the following scale (write down the number in the table):

Rarely Sometimes Frequently Very Frequently

1 _________ 2____________3_________ 4___________

Relationship	Languages mixed	Frequency of mixing
Spouse/family members		
Friends		
Co-workers		
Classmates		

20. Specify which language is typically used under each condition:

	At home	At work
Reading	___________	___________
Writing	___________	___________
Speaking	___________	___________

21. Among the languages you know, which language is the one that you would prefer to use in these situations?

At home	___________	At a party	___________
At work	___________	In general	___________

22. If you have taken a standardized test of proficiency for your second language (e.g., TOEFL or Test of English as a Foreign Language), please indicate the scores you received for each.

Language	**Writing**	**Speaking**	**Reading**	**Listening**	**Name of the Test**
________	______	_______	________	_______	________

23. If there is anything else that you feel is interesting or important about your language background or language use, please comment below.

(Adapted from www.personal.psu.edu*)*

APPENDIX D

Word and nonword stimuli used in the lexical decision task

Item	Type	Meaning	Relatedness	Letter	Frequency	Neighbor	Number of Meanings	Relatedness ratings
mine	word	few	low	4	59	27	4	2.98
watch	word	few	low	5	81	9	4	1095
firm	word	few	low	4	109	5	4	1.84
fast	word	few	low	4	78	13	3	2.48
rich	word	few	low	4	74	5	4	2.63
dull	word	few	high	4	27	15	4	4.21
ship	word	few	high	4	83	11	2	4.21
dust	word	few	high	4	70	12	3	4.26
drink	word	few	high	5	82	5	3	5.04
ring	word	many	low	4	47	13	9	2.84
scale	word	many	low	5	60	7	9	2.79
check	word	many	low	5	88	4	9	2.08
pitch	word	many	low	5	22	9	7	2.45
cross	word	many	low	5	55	6	6	2.09
blank	word	many	high	5	14	7	6	3.81
smoke	word	many	high	5	41	4	8	3.81
clean	word	many	high	5	70	3	7	5.2
file	word	many	high	4	81	16	8	3.71
share	word	many	high	5	98	16	6	4.51
dump	word	many	high	4	4	9	6	3.94
bill	word	many	high	4	143	17	6	3.52
pound	word	few	low	5	28	7	4	1.54
chest	word	few	low	5	53	4	3	2.11
trip	word	few	low	4	81	8	4	2.00
date	word	few	low	4	103	221	4	2.22
hide	word	few	low	4	22	13	2	1.64
park	word	few	low	4	94	15	2	1.76
ball	word	few	low	4	110	20	4	1.57
bomb	word	few	high	4	36	4	4	4.32
story	word	few	high	5	153	4	3	3.63
card	word	few	high	4	26	14	3	3.55

Item	Type	Meaning	Relatedness	Letter	Frequency	Neighbor	Number of Meanings	Relatedness ratings
shop	word	few	high	4	63	12	3	3.96
rule	word	few	high	4	73	9	3	3.71
land	word	few	high	4	217	15	4	3.67
strip	word	many	low	5	30	4	7	2.15
cast	word	many	low	4	45	19	8	2.12
draw	word	many	low	4	56	6	10	1.32
sharp	word	many	low	5	72	5	7	2.89
stock	word	many	low	5	147	6	7	2.24
plot	word	many	low	4	37	8	6	2.93
club	word	many	low	4	145	2	6	1.9
lock	word	many	high	4	23	16	6	4.16
bound	word	many	high	5	42	7	9	3.68
round	word	many	high	5	81	7	13	3.51
cover	word	many	high	5	88	13	10	4.09
fine	word	many	high	4	161	23	9	3.53
limit	word	few	high	5	48	0	3	5.84
safe	word	few	high	4	58	8	3	5.79
beaf	nonword							
korn	nonword							
durt	nonword							
doar	nonword							
feer	nonword							
frum	nonword							
heet	nonword							
jeap	nonword							
lyne	nonword							
lyke	nonword							
neet	nonword							
roze	nonword							
scid	nonword							
scin	nonword							
soop	nonword							
tyme	nonword							
dreem	nonword							
furst	nonword							

Item	Type	Meaning	Relatedness	Letter	Frequency	Neighbor	Number of Meanings	Relatedness ratings
aftur	nonword							
elboe	nonword							
sneek	nonword							
nurve	nonword							
wheet	nonword							
burth	nonword							
treet	nonword							
swerl	nonword							
sheap	nonword							
speer	nonword							
majic	nonword							
muzic	nonword							
wheal	nonword							
berch	nonword							
proze	nonword							
speek	nonword							
urbin	nonword							
phaze	nonword							
leest	nonword							
durby	nonword							
panzy	nonword							
sheat	nonword							
leef	nonword							
teem	nonword							
hert	nonword							
ferm	nonword							
amuze	nonword							
skar	nonword							
bernt	nonword							
kare	nonword							

Notes. The data was extracted from Azuma and Van Orden's list (1997).

APPENDIX E

Picture Stimuli and Prime Words Used in Experiment one (chapter four)

i. Characteristics of cognate and non-cognate pictures (experiment one)

Cognate Pictures	Syll	Ch	Freq	AOA	P	Non-Cognate picture	Syll	Ch	Freq	AOA	P
balcony	3	7	2.639	3	3.42	lettuce	2	7	2.079	3	2.4
balloon	2	7	1.946	1	4.5	eagle	2	5	2.3026	3	2.1
bus	1	3	4.382	1	5	egg	1	3	4.4659	1	1.23
cake	1	4	3.555	1	5	leaf	1	4	4.407	3	2.5
camel	2	5	3.258	3	3.87	bottle	2	6	4.7622	1	1.7
camera	3	6	3.611	2	4.32	cheese	1	6	3.4657	1	1.18
cigarette	3	9	4.277	3	3.81	envelope	3	8	3.2189	3	1.35
dinosaur	3	8	1.792	3	4.25	tiger	2	5	2.565	1	1.21
doctor	2	6	5.220	2	5	chicken	2	7	3.7377	1	1.29
dolphin	2	7	1.386	3	5	ostrich	2	7	1.386	3	1.02
giraffe	2	7	1.099	1	4.56	pencil	2	6	2.9957	2	2.1
gorilla	3	7	1.386	3	4.37	pyramid	3	7	2.079	3	1.1
jacket	2	6	3.761	1	4.02	rabbit	2	6	2.9957	3	1.08
jar	1	3	2.996	2	3.9	fox	1	3	2.7726	3	1.5
kangaroo	3	8	1.386	3	4.15	bicycle	3	7	1.7918	1	1.34
lamp	1	4	3.584	1	5	skirt	1	5	3.401	3	1.77
lemon	2	5	2.773	3	5	pillow	2	6	2.9957	1	1.25
music	2	5	4.898	3	3.7	flower	2	6	4.5433	1	1.06

Cognate Pictures	Syll	Ch	Freq	AOA	P	Non-Cognate picture	Syll	Ch	Freq	AOA	P
octopus	3	7	1.099	3	4.3	penguin	2	7	1.792	1	1.12
piano	3	5	3.332	3	5	wallet	2	6	2.1972	3	2.1
pizza	2	5	1.099	1	4.8	turtle	2	6	1.609	1	1.5
potato	3	6	3.611	2	4.3	needle	2	6	2.8332	3	1.09
radio	3	5	4.489	1	5	cat	1	3	4.22	1	1.8
telephone	3	9	4.663	1	3.71	elephant	3	8	3.2189	1	1.22
tv	2	2	0.000	1	3.22	ring	1	4	1.386	3	1.4
tomato	3	6	2.708	3	3.91	candle	2	6	2.8332	3	1.77
genie	2	5	0.693	3	5	saw	1	3	0.693	3	1
panda	2	5	0.693	3	4.13	grapes	1	6	0	2	1.46
sandwich	2	8	0.000	2	4.68	curtains	2	8	0	3	1.02
helicopter	4	10	2.833	2	4.55	fountain	2	8	2.5649	3	1.8

Notes. Syll refers to the number of syllables, Ch refers to number of characters, Freq refers to mean frequency count, and AOA refers to the age of acquisition. These data were extracted from the International Picture Naming Project online-database (Szekely et al. 2003; Szekely et al. 2004). P refers to the mean phonological similarity rating obtained from the rating task performed by Arabic-English bilinguals as detailed in chapter four.

ii. **Characteristics of prime words for cognate and non-cognate pictures (experiment one)**

Picture	Type	L1 Translation	Syll	Ch	Fre	L1 Prime Unrelated	Meaning	Syll	Ch	Fre
balcony	cognate	بلكونه	4	5	0.8	فأر	mouse	2	3	0.7
balloon	cognate	بالونه	3	6	0.23	مشط	comb	1	4	0.21
bus	cognate	باص	1	4	7.78	بطريق	penguin	2	5	8.15
cake	cognate	كيك	2	5	0.87	شرطي	policewoman	2	4	0.73
camel	cognate	جمل	2	5	2.94	ريشة	feather	2	4	2.52
camera	cognate	كاميرا	3	5	4.40	بصله	onion	2	5	3.90
cigarette	cognate	سيجارة	3	6	3.98	حزام	belt	2	5	3.9
dinosaur	cognate	ديناصور	4	7	0.34	عنكبوت	spider	3	8	0.34
doctor	cognate	دكتور	3	7	0.49	مروحة	fan	3	7	0.46
dolphin	cognate	دولفين	2	6	0.56	اسد	lion	2	5	0.6
giraffe	cognate	زرافة	3	6	0.12	ولاعة	lighter	3	6	0.13
gorilla	cognate	غوريلا	3	6	0.32	مكنسة	broom	3	6	0.29
jacket	cognate	جاكيت	2	5	1.59	عظمة	bone	3	5	1.69
jar	cognate	جرة	2	3	0.18	تمساح	crocodile	2	6	0.16
kangaroo	cognate	كنغر	2	4	0.23	نملة	ant	2	5	0.23
lamp	cognate	لمبة	3	6	1.45	ساحرة	witch	3	5	1.66
lemon	cognate	ليمون	2	5	0.6	عقرب	scorpian	2	6	0.6
music	cognate	موسيقى	2	5	4.55	ملاك	angel	2	5	4.55
octopus	cognate	اخطبوط	3	5	0.86	خروف	sheep	3	5	0.78
piano	cognate	بيانو	2	5	0.66	حمل	lamb	2	5	0.68

Picture	Type	L1 Translation	Syll	Ch	Fre	L1 Prime Unrelated	Meaning	Syll	Ch	Fre
pizza	cognate	بيتزا	3	5	0.19	فاس	axe	1	3	0.21
potato	cognate	بطاطا	2	6	6.14	قلم	pen	2	5	5.33
radio	cognate	راديوا	2	6	3.7	نحلة	bee	2	5	3.02
telephone	cognate	تليفون	2	5	17.19	رسالة	letter	3	6	16.41
television	cognate	تلفاز	2	5	0.05	ذرة	corn	2	4	0.05
tomato	cognate	طماطم	2	5	0.26	سخان	heater	3	6	0.26
genie	cognate	جني	3	4	0.13	بومة	owl	2	4	0.11
panda	cognate	باندا	2	5	0.29	سنجاب	Squirrel	2	6	0.26
sandwich	cognate	ساندوتش	3	6	9.3	سطل	pocket	2	5	9.94
helicopter	cognate	هيلكبتر	4	7	0.23	كورة	ball	3	5	0.21
bicycle	Non-cog	دراجة	3	5	2.78	مسدس	gun	3	4	2.73
bottle	Non-cog	قارورة	3	5	2.89	دب	bear	1	3	2.16
candle	Non-cog	شمعة	2	4	2.52	صور	fence	1	3	2.57
leaf	Non-cog	ورقة	3	4	0.51	بطة	duck	2	4	0.49
cheese	Non-cog	جبنة	2	5	0.55	ملعقه	spoon	2	5	0.52
chicken	Non-cog	دجاجة	3	5	0.75	فطر	mushroom	2	5	0.73
curtains	Non-cog	ستارة	3	5	1.09	عروسة	doll	3	5	1.22
eagle	Non-cog	صقر	1	3	11.99	شباك	window	2	4	10.96
egg	Non-cog	بيض	1	3	2.55	كفر	tire	2	3	3.02
elephant	Non-cog	فيل	1	3	2.22	زرار	button	1	3	1.92
envelope	Non-cog	ظرف	1	3	8.27	كرسي	chair	2	4	8.22
pencil	Non-cog	مرسام	3	5	5.33	قارب	boat	3	4	5.67
pillow	Non-cog	مخدة	3	4	1.33	ثلاجة	frideg	3	5	1.14
rabbit	Non-cog	ارنب	2	4	0.05	قرعة	pumpkin	2	4	0.03

Picture	Type	L1 Translation	Syll	Ch	Fre	L1 PrimeUnrelated	Meaning	Syll	Ch	Fre
wallet	Non-cog	محفظة	3	3	3.64	ذراع	arm	3	3	3.68
needle	Non-cog	ابرة	2	4	0.53	موزة	banana	2	4	0.42
penguin	Non-cog	بطريق	2	5	8.15	سلة	basket	2	4	8.50
skirt	Non-cog	تنورة	3	5	1.35	حمار	donkey	3	4	1.69
fox	Non-cog	ثعلب	2	4	2.21	جرس	bell	2	3	2.16
ring	Non-cog	خاتم	2	5	2.18	ابريق	teapot	2	5	2.12
flower	Non-cog	زهرة	2	4	5.20	حصان	horse	1	3	4.59
turtle	Non-cog	سلحفاة	3	7	0.39	مسطرة	ruler	3	7	0.39
grapes	Non-cog	عنب	2	3	0.78	معلمة	teacher	4	5	0.70
cat	Non-cog	قطة	2	3	0.96	ربطة	tie	2	3	0.94
lettuce	Non-cog	خس	1	3	0.10	جيوب	bucket	2	4	0.08
saw	Non-cog	منشار	2	4	0.03	تاج	crown	1	3	0.03
fountain	Non-cog	نافورة	3	5	5.20	عصفور	bird	3	5	5.62
ostrich	Non-cog	نعامة	4	6	0.39	سبحة	rosary	3	6	0.42
tiger	Non-cog	نمر	1	5	12.43	قطار	train	2	5	12.25
pyramid	Non-cog	اهرامات	3	7	0.43	برتقالة	orange	3	6	0.42

Notes. Syll refers to the number of syllables, Ch refers to number of characters, and Freq refers to mean frequency count. These data were extracted from the "The Gulf Arabic Nouns" (Khwaileh et al. 2018) and the lexical database for Modern Standard Arabic "ARALEX" (Boudelaa and Marslen-Wilson 2010).

iii. Characteristics of filler items (experiment one)

Picture Name	Syllable length	Character length	Frequency	Age of acquisition	Unrelated primes	Syllable length	character length	Frequency
waiter	2	6	3.135	3	worm	2	4	1.07
pen	1	3	3.296	1	baby	2	5	23.02
crab	3	6	1.61	3	drawer	1	4	3.69
fish	1	4	5.100	1	slide	3	6	7.54
peach	1	5	1.946	3	ladder	2	5	29.05
fork	1	4	2.773	1	bag	2	5	0.44
mosquito	3	8	1.792	3	dates	1	5	3.43
clown	1	5	1.609	2	flag	2	5	94.23
turkey	2	6	1.792	1	kite	3	5	25.63
mountain	1	4	75.06	3	box	2	6	89.68
fireman	3	7	1.609	2	thread	1	3	2.86
bug	1	3	3.761	1	hat	3	7	0.55
zebra	2	5	1.099	2	rope	2	5	1.82
clock	1	5	3.689	1	sun	1	4	75.06
asparagus	4	9	1.099	3	bull	1	3	0.88
mirror	2	6	3.912	3	key	2	6	13.76
sailboat	2	8	0.000	3	slipper	2	4	0.03
mixer	2	5	1.099	3	truck	2	5	7.71
nail	1	4	3.258	2	bridge	2	5	14.30
cow	1	4	0.34	1	curtain	3	5	1.09

Note. These characteristics of picture names were extracted from the International Picture Naming Project online-database (Szekely et al. 2003; Szekely et al. 2004). The characteristics of prime words (syllable length, character length and frequency are taken from the "The Gulf Arabic Nouns" (Khwaileh et al. 2018) and the lexical database for Modern Standard Arabic "ARALEX" (Boudelaa and Marslen-Wilson 2010).

APPENDIX F

Picture stimuli and prime words used in Experiment two (chapter five)

i. Characteristics of non-cognate pictures (experiment two)

Picture	Syllable Length	Character length	Frequency	Age of Acquisition
cow	1	3	3.714	1
girl	1	4	6.084	1
frog	1	4	2.303	1
horse	1	5	4.890	1
snake	1	5	3.178	3
nurse	1	5	3.912	2
pig	1	3	3.784	1
tiger	2	5	2.565	1
dog	1	3	4.754	1
chicken	2	7	3.738	1
monkey	2	6	2.944	1
fish	1	4	5.100	1
bat	1	3	2.708	2
fork	1	4	2.773	1
carrot	2	6	2.197	1
bed	1	3	5.136	1
car	1	3	5.872	1
door	1	4	5.958	1
pillow	2	6	2.996	1
plate	1	5	4.025	1
drum	1	4	2.833	3

Picture	Syllable Length	Character length	Frequency	Age of Acquisition
deer	1	4	2.565	1
butterfly	3	9	2.398	1
fireman	3	7	1.609	2
wolf	1	4	2.398	2
mosquito	3	8	1.792	3
apple	2	5	3.434	1
strawberry	3	10	1.946	2
hammer	2	6	2.485	1
scarf	1	5	2.565	2

Note. These data were extracted from the International Picture Naming Project online-database (Szekely et al. 2003; Szekely et al. 2004).

ii. Characteristics of L1 prime words (experiment two)

Picture	Related Prime	Meaning	Syll	Ch	Fre	Semantic similarity Rating	Unrelated Prime	Meaning	Syll	Ch	Fre	Semantic similarity Rating
cow	ماعز	goat	1	4	0.34	4.52	سماعة	headphone s	2	5	0.23	1.3
girl	صبي	boy	2	5	167.80	4.38	كتاب	book	2	5	152.85	1.74
frog	سلحفاة	turtle	3	7	0.39	4.32	مسطرة	ruler	3	7	0.39	1.12
horse	حمار	donkey	3	4	1.69	4	عظمة	bone	3	5	1.69	1.04
snake	تمساح	crocodile	2	6	0.16	4.24	مرساة	anchor	2	5	0.16	1
nurse	معلمة	teacher	4	5	0.70	4.42	جبنه	cheese	2	5	0.55	1.14
pig	خروف	sheep	3	5	0.68	4.56	طاقية	hat	3	7	0.55	1.44
tiger	اسد	lion	2	5	0.6	4.58	سطل	bucket	2	4	0.08	1.1
dog	قطة	cat	2	3	0.96	4.48	مقبض	doorknob	2	4	0.81	1.2
chicken	بطة	duck	2	4	0.49	4.56	حقيبة	bag	2	5	0.44	1.44
monkey	فهد	cheetah	2	5	64.22	4.3	عين	eye	1	3	64.06	1.22
fish	بطريق	penguin	2	5	8.15	4.46	سلة	basket	2	4	8.50	1.76
bat	فأر	mouse	1	3	0.7	3.82	ولاعة	lighter	3	6	0.13	1.26
fork	ملعقه	spoon	2	5	0.52	4.38	قرد	monkey	2	5	0.47	1.92
carrot	بصل	onion	2	5	3.90	4.48	حزام	belt	2	5	3.9	1.9
bed	كرسي	chair	2	4	8.22	4.56	زحليقة	slide	3	6	7.54	1.84
car	قطار	train	2	5	12.25	3.76	صقر	falcon	1	3	11.99	1.18
door	شباك	window	2	4	10.96	4.54	كفر	tire	2	3	3.02	1.42
pillow	درج	drawer	1	4	3.69	4.32	ذراع	arm	3	3	3.68	1.78
plate	ابريق	teapot	2	5	2.12	4.56	دب	bear	1	3	2.16	1.78

Picture	Related Prime	Meaning	Syll	Ch	Fre	Semantic similarity Rating	Unrelated Prime	Meaning	Syll	Ch	Fre	Semantic similarity Rating
drum	ناي	flute	2	6	0.18	4.48	بومة	owl	2	4	0.11	1.94
deer	حمل	lamb	2	5	0.68	4.84	كبوس	cap	2	5	0.68	1
butterfly	عنكبوت	spider	3	8	0.34	3.88	مكنسة	broom	3	6	0.29	1.34
fireman	شرطي	policewo man	2	4	0.73	4.72	ورقة	leaf	3	4	0.51	1.4
wolf	ثعلب	fox	2	4	2.21	4.3	شمعة	candle	2	4	2.52	1.38
mosquito	دودة	worm	2	4	1.07	4.74	حبل	rope	2	5	1.82	1.24
apple	برتقال	orange	3	6	0.42	3.96	مسبح	rosary	3	6	0.42	1
strawber ry	موزه	banana	2	4	0.42	3.64	مروحة	fan	3	7	0.46	1.34
hammer	منشار	axe	1	3	0.21	4.2	كورة	ball	3	5	0.21	1.6
scarf	ربطة	tie	2	3	0.94	4.44	ثور	bull	1	3	0.88	1.58

Notes. Syll refers to the number of syllables, Ch refers to number of characters, and Freq refers to mean frequency count. These data were extracted from the “The Gulf Arabic Nouns” (Khwaileh et al. 2018) and the lexical database for Modern Standard Arabic “ARALEX” (Boudelaa and Marslen-Wilson 2010). The semantic similarity rating refers to the mean rating of similarity and is taken from the similarity rating task reported in chapter five.

iii. Characteristics of filler items (experiment two)

Picture Name	Syllable Length	Character length	Frequency	Age of acquisition	Unrelated Primes	Syllable Length	Character length	Frequency
waiter	2	6	3.135	3	Ostrich	4	6	0.39
pen	1	3	3.296	1	baby	2	5	23.02
crab	3	6	1.61	3	fence	1	3	2.57
skirt	1	5	3.401	3	angel	2	5	4.55
peach	1	5	1.946	3	ladder	2	5	29.05
rabbit	2	6	0.26	3	wallet	1	5	3.64
lettuce	2	5	0.73	3	dates	1	5	3.43
clown	1	5	1.609	2	flag	2	5	94.23
turkey	2	6	1.792	1	kite	3	5	25.63
mountain	1	4	75.06	3	box	2	6	89.68
pumpkin	2	7	1.099	2	thread	1	3	2.86
bug	1	3	3.761	1	grapes	2	5	0.78
zebra	2	5	1.099	2	egg	2	4	2.55
clock	1	5	3.689	1	sun	1	4	75.06
asparagus	4	9	1.099	3	bottle	2	5	0.21
mirror	2	6	3.912	3	key	2	6	13.76
sailboat	2	8	0.000	3	slipper	2	4	0.03
mixer	2	5	1.099	3	truck	2	5	7.71
nail	1	4	3.258	2	bridge	2	5	14.30
ant	1	3	2.565	2	curtain	3	5	1.09

Note. These characteristics of picture names were extracted from the International Picture Naming Project online-database (Szekely et al. 2003; Szekely et al. 2004). The characteristics of prime words (syllable length, character length and frequency are taken from the "The Gulf Arabic Nouns" (Khwaileh et al. 2018) and the lexical database for Modern Standard Arabic "ARALEX" (Boudelaa and Marslen-Wilson 2010).

APPENDIX G

Picture Stimuli and Prime Words Used in Experiment three (chapter six)

i. Characteristics of non-cognate pictures (experiment three)

Pictures	Type	Syllable Length	Character length	Frequency	Age of Acquisition
cow	Animate	1	3	3.714	1
girl	Animate	1	4	6.084	1
parrot	Animate	2	6	1.609	3
butterfly	Animate	3	9	2.398	1
crab	Animate	1	4	2.303	3
frog	Animate	1	4	2.303	1
horse	Animate	1	5	4.890	1
snake	Animate	1	5	3.178	3
nurse	Animate	1	5	3.912	2
pig	Animate	1	3	3.784	1
tiger	Animate	2	5	2.565	1
dog	Animate	1	3	4.754	1
chicken	Animate	2	7	3.738	1
fly	Animate	1	3	3.611	3
elephant	Animate	3	8	3.219	1
rabbit	Animate	2	6	2.996	3
deer	Animate	1	4	2.565	1
king	Animate	1	4	4.605	3
monkey	Animate	2	6	2.944	1
pirate	Animate	2	6	1.792	3
motorbike	Inanimate	4	10	2.708	1

Pictures	Type	Syllable Length	Character length	Frequency	Age of Acquisition
apple	Inanimate	2	5	3.434	1
strawberry	Inanimate	3	10	1.946	2
lettuce	Inanimate	2	7	2.079	3
carrot	Inanimate	2	6	2.197	1
leaf	Inanimate	1	4	4.407	3
pineapple	Inanimate	3	9	1.386	3
mountain	Inanimate	2	8	4.443	3
bed	Inanimate	1	3	5.136	1
car	Inanimate	1	3	5.872	1
hammer	Inanimate	2	6	2.485	1
scarf	Inanimate	1	5	2.565	2
drum	Inanimate	1	4	2.833	3
door	Inanimate	1	4	5.958	1
helmet	Inanimate	2	6	2.639	3
stove	Inanimate	1	5	3.045	1
castle	Inanimate	2	6	3.332	3
pillow	Inanimate	2	6	2.996	1
washing machine	Inanimate	4	14	0.693	3
plate	Inanimate	1	5	4.025	1

Notes. These data were extracted from the International Picture Naming Project online-database (Szekely et al. 2003; Szekely et al. 2004).

ii. Characteristics of L1 primes (experiment three)

Picture	Related Prime	Meaning	Syll	Ch	Fre	Semantic similarity Rating	Unrelated Prime	Meaning	Syll	Ch	Fre	Semantic similarity Rating
cow	غنمة	goat	1	4	0.34	4.41	شرطي	policewoman	2	4	0.73	1.26
girl	ولد	boy	2	5	167.80	4.32	كتاب	book	2	5	152.85	1.54
parrot	بومة	owl	2	4	0.11	4.35	سماعة	headphones	2	5	0.23	1.08
butterfly	عنكبوت	spider	3	8	0.34	3.96	مسطرة	ruler	3	7	0.39	1.29
crab	صدفه	seashell	3	6	1.61	4.41	ستارة	curtain	3	5	1.09	1.10
frog	سلحفاة	turtle	3	7	0.39	4.38	شنطة	bag	2	5	0.44	1.15
horse	حمار	donkey	2	4	1.69	4.53	مغني	singer	3	7	1.90	1.33
snake	تمساح	crocodile	2	6	0.16	4.44	جره	vase	2	4	0.18	1.35
nurse	معلمة	teacher	4	5	0.70	4.56	قرد	monkey	2	5	0.47	1.19
pig	خروف	sheep	3	5	0.68	4.53	طاقيه	hat	3	7	0.55	1.28
tiger	اسد	lion	2	5	0.6	4.44	سطل	bucket	2	4	0.08	1.36
dog	قطة	cat	2	3	0.96	4.44	دوده	worm	2	4	1.07	1.49
chicken	بطه	duck	2	4	0.49	3.47	مروحة	fan	3	7	0.46	1.32
fly	نحلة	bee	2	5	3.02	4.41	بصل	onion	2	5	3.90	1.62
elephant	وحيد القرن	Rhino	5	9	16.13	4.47	جسر	bridge	2	5	14.30	1.91
rabbit	سنجاب	Squirrel	2	6	0.26	4.45	ولاعه	lighter	3	6	0.13	1.64
deer	حمل	lamb	2	5	0.68	3.71	جبنة	cheese	2	5	0.55	1.13
king	ساحرة	witch	2	5	1.66	4.62	حبل	rope	2	5	1.82	1.32
monkey	نمر	cheetah	2	5	64.22	4.35	قلب	heart	1	4	64.81	1.63
pirate	مغني	singer	3	7	1.90	4.56	شمعة	candle	2	4	2.52	1.78
motorbike	دراجة	bicycle	2	5	2.78	4.45	دب	bear	1	2	2.16	1.71
apple	برتقال	orange	3	6	0.42	4.65	سبحة	rosary	3	6	0.42	1.06
strawberry	موزة	banana	2	4	0.42	3.85	نعامه	Ostrich	4	6	0.39	1.22
lettuce	فطر	mushroom	2	5	0.73	4.46	معلمة	teacher	4	5	0.70	1.29

Picture	Related Prime	Meaning	Syll	Ch	Fre	Semantic similarity Rating	Unrelated Prime	Meaning	Syll	Ch	Fre	Semantic similarity Rating
carrot	بصل	onion	2	5	3.90	4.36	كفر	tire	2	3	3.02	1.42
leaf	جذور	roots	2	5	13.76	4.67	شاحنة	truck	2	5	7.71	1.17
pineapple	عنب	grapes	2	5	0.78	3.76	كبوس	cap	2	5	0.68	1.00
mountain	شمس	sun	1	4	75.06	3.49	عين	eye	1	3	64.06	1.36
bed	كرسي	chair	2	4	8.22	4.38	بطريق	penguin	2	5	8.15	1.59
car	قطار	train	2	5	12.25	4.31	صقر	falcon	1	3	11.99	1.46
hammer	فاس	axe	1	3	0.21	3.73	شبشب	slipper	2	4	0.03	1.29
scarf	ربط عنق	tie	2	3	0.94	4.51	مقبض	doorknob	2	4	0.81	1.72
drum	مزمار	flute	2	6	0.18	3.26	عنكبوت	spider	3	8	0.34	1.09
door	شباك	window	2	4	10.96	4.35	سلة	basket	2	4	8.50	1.63
helmet	تاج	crown	1	3	0.03	4.29	كمثرى	pear	3	7	0.03	1.35
stove	سخان	heater	3	6	0.26	3.55	قارب	boat	3	6	5.67	1.37
castle	كنيسة	church	3	6	18.41	4.29	مفتاح	key	2	6	13.76	1.22
pillow	درج	drawer	1	4	3.69	4.71	بيض	egg	2	4	2.55	1.74
washing machine	ثلاجة	frideg	3	6	1.14	3.76	عظمة	bone	3	5	1.69	1.51
plate	برادي	teapot	2	5	2.12	4.56	نحلة	bee	2	5	3.02	1.28

Notes. Syll refers to the number of syllables, Ch refers to number of characters, and Freq refers to mean frequency count. These data were extracted from the "The Gulf Arabic Nouns" (Khwaileh et al. 2018) and the lexical database for Modern Standard Arabic "ARALEX" (Boudelaa and Marslen-Wilson 2010). The semantic similarity rating refers to the mean rating of similarity and is taken from the similarity rating task reported in chapter six.

iii. Characteristics of filler items (experiment three)

Picture Name	Syllable Length	Character length	Frequency	Age of acquisition	Unrelated Primes	Syllable Length	Character length	Frequency
waiter	2	6	3.135	3	ball	3	5	2.89
pen	1	3	3.296	1	baby	2	5	23.02
fish	1	4	5.100	1	fence	1	3	2.57
skirt	1	5	3.401	3	angel	2	5	4.55
peach	1	5	1.946	3	ladder	2	5	29.05
mosquito	3	8	1.792	3	wallet	1	5	3.64
fireman	3	7	1.609	2	dates	1	5	3.43
clown	1	5	1.609	2	flag	2	5	94.23
turkey	2	6	1.792	1	kite	3	5	25.63
fork	1	4	2.773	1	box	2	6	89.68
pumpkin	2	7	1.099	2	thread	1	3	2.86
bug	1	3	3.761	1	belt	2	5	3.9
zebra	2	5	1.099	2	slide	3	6	7.54
clock	1	5	3.689	1	arm	3	3	3.68
asparagus	4	9	1.099	3	bottle	2	5	0.21
mirror	2	6	3.912	3	bull	1	3	0.88
sailboat	2	8	0.000	3	broom	3	6	0.29
mixer	2	5	1.099	3	fox	2	4	2.21
nail	1	4	3.258	2	mouse	1	3	0.7
ant	1	3	2.565	2	spoon	2	5	0.52

Note. These characteristics of picture names were extracted from the International Picture Naming Project online-database (Szekely et al. 2003; Szekely et al. 2004). The characteristics of prime words (syllable length, character length and frequency are taken from the "The Gulf Arabic Nouns" (Khwaileh et al. 2018) and the lexical database for Modern Standard Arabic "ARALEX" (Boudelaa and Marslen-Wilson 2010).

APPENDIX H

Picture Stimuli and Phonemes Used in Experiment four (chapter seven)

i. Characteristics of picture stimuli (experiment four)

Picture	Syllable Length	Character length	Frequency	Age of Acquisition
ball	1	4	4.718	1
ladder	2	6	2.833	2
teeth	1	5	1.386	3
dog	1	3	4.754	1
leaf	1	4	4.407	3
well	1	4	1.792	3
witch	1	5	3.497	3
hammer	2	6	2.485	1
hand	1	4	6.586	1
door	1	4	5.958	1
elephant	3	8	3.219	1
bridge	1	6	4.205	3
rabbit	2	6	2.996	3
fan	1	3	2.890	3
strawberry	3	10	1.946	2

Picture	Syllable Length	Character length	Frequency	Age of Acquisition
book	1	4	6.075	1
house	1	5	6.409	1
king	1	4	4.605	3
razor	2	5	2.303	3
butterfly	3	9	2.398	1
feather	2	7	3.091	3
needle	2	6	2.833	3
bed	1	3	5.136	1
turtle	2	6	1.609	1
banana	3	6	2.197	1
skirt	1	5	3.401	3
leg	1	3	5.170	1
stairs	1	6	3.807	1
tire	2	4	2.485	3
fire	2	4	5.094	3
basket	2	6	3.219	2
foot	1	4	5.790	1
toothbrush	2	10	1.099	1
lion	2	4	3.258	1
bear	1	4	2.833	1
tiger	2	5	2.565	1

Picture	Syllable Length	Character length	Frequency	Age of Acquisition
mountain	2	8	4.443	3
scarf	1	5	2.565	2
dress	1	5	4.477	1
flower	2	6	4.534	1
fish	1	4	5.100	1
bell	1	4	3.332	3
mouse	1	5	2.944	1
cat	1	3	4.220	1
duck	1	4	1.230	1

Notes. These data were extracted from the International Picture Naming Project online-database (Szekely et al. 2003; Szekely et al. 2004).

ii. The target phonemes used in the three experimental conditions (experiment four)

Picture Name	Arabic Name	Control Phonemes	English Phonemes	Arabic Phonemes
duck	بطة	/t/	/d/	/b/
cat	بسة	/ʒ/	/k/	/b/
basket	سلة	/r/	/b/	/s/
bear	دب	/ʒ/	/b/	/d/
bell	جرس	/f/	/b/	/ʒ/
dress	فستان	/ʒ/	/d/	/f/
fish	سمكة	/r/	/f/	/s/
flower	وردة	/j/	/f/	/w/
foot	رجل	/j/	/f/	/l/
lion	اسد	/ʒ/	/l/	/s/
mountain	جبل	/r/	/m/	/ʒ/
mouse	فأر	/d/	/m/	/f/
scarf	وشاح	/n/	/s/	/w/
tiger	نمر	/w/	/t/	/n/
toothbrush	فرشاة	/n/	/t/	/f/
banana	موز	/r/	/b/	/m/
bed	سرير	/l/	/d/	/s/
book	كتاب	/l/	/k/	/k/
butterfly	فراشة	/n/	/b/	/f/
feather	ريشة	/ʒ/	/f/	/r/
fire	نار	/h/	/f/	/n/
house	بيت	/k/	/h/	/b/

Picture Name	Arabic Name	Control Phonemes	English Phonemes	Arabic Phonemes
king	ملك	/w/	/k/	/m/
leg	ساق	/n/	/l/	/r/
needle	ابرة	/w/	/n/	/b/
razor	موس	/r/	/r/	/m/
skirt	تنورة	/d/	/s/	/t/
stairs	درج	/n/	/s/	/d/
turtle	سلحفاه	/d/	/t/	/s/
tyre	كفر	/ʒ/	/t/	/k/
ball	كرة	/ʒ/	/b/	/k/
bridge	جسر	/t/	/b/	/ʒ/
dog	كلب	/m/	/d/	/k/
door	باب	/j/	/d/	/b/
elephant	فيل	/d/	/l/	/f/
fan	مروحة	/j/	/f/	/m/
hammer	مطرقة	/k/	/h/	/m/
hand	يد	/r/	/h/	/j/
ladder	سلم	/h/	/l/	/s/
leaf	ورقة	/n/	/l/	/w/
rabbit	ارنب	/m/	/b/	/r/
strawberry	فراولة	/ʒ/	/s/	/f/
teeth	اسنان	/d/	/t/	/s/
well	بئر	/k/	/w/	/b/
witch	ساحرة	/l/	/w/	/s/

iii. Filler pictures (experiment four)

Picture	Arabic Name	Filler	Syllable Length	character length	frequency	age of acquisition
cow	بقرة	/s/	1	3	3.714	1
girl	بنت	/f/	1	4	6.084	1
frog	ضفدع	/j/	1	4	2.303	1
horse	حصان	/m/	1	5	4.890	1
snake	حيه	/t/	1	5	3.178	3
nurse	ممرضه	/w/	1	5	3.912	2
pig	خنزير	/h/	1	3	3.784	1
chicken	دجاج	/l/	2	7	3.738	1
monkey	قرد	/t/	2	6	2.944	1
bat	خفاش	/l/	1	3	2.708	2
fork	شوكة	/j/	1	4	2.773	1
carrot	جزر	/h/	2	6	2.197	1
car	سيارة	/w/	1	3	5.872	1
pillow	مخدة	/s/	2	6	2.996	1
plate	صحن	/h/	1	5	4.025	1
drum	طبلة	/k/	1	4	2.833	3
deer	غزال	/f/	1	4	2.565	1
fireman	رجل اطفاء	/h/	3	7	1.609	2
wolf	ذئب	/f/	1	4	2.398	2
apple	تفاح	/k/	2	5	3.434	1

Notes. The data (i.e., no of syllable and characters, frequency rating and age of acquisition) were extracted from the International Picture Naming Project online-database (Szekely et al. 2003; Szekely et al. 2004).

APPENDIX I

Picture Stimuli and Phonemes Used in Experiment five (chapter eight)

i. Characteristics of picture stimuli (experiment five)

Picture Name	Syllable Length	Character length	Frequency	Age of Acquisition
picture	2	7	5.165	1
hair	1	4	5.298	3
donkey	2	6	2.708	1
hat	1	3	4.234	1
lettuce	2	7	2.079	3
onion	2	5	2.833	3
nail	1	4	3.258	2
grapes	1	6	0.000	2
duck	1	4	1.230	1
parrot	2	6	1.609	3
bird	1	4	0.000	1
hanger	2	6	1.099	3
diaper	2	6	1.099	1
drum	1	4	2.833	3
shell	1	5	3.850	3
frog	1	4	2.303	1
table	2	5	3.464	1
horse	1	5	4.890	1

Picture Name	Syllable Length	Character length	Frequency	Age of Acquisition
plate	1	5	4.025	1
cloud	1	5	4.043	2
envelope	3	8	3.219	3
flag	1	4	3.296	2
hammer	2	6	2.485	1
ring	1	4	1.386	3
belt	1	4	3.296	2
scissors	2	8	1.609	1
bone	1	4	4.248	3
bread	1	5	4.317	1
submarine	3	9	2.890	3
bride	1	5	2.565	3
desert	2	6	3.738	3
doll	1	4	3.258	1
eagle	2	5	2.303	3
earring	2	7	1.386	3
finger	2	6	4.820	1
goat	1	4	3.367	3
mosquito	3	8	1.792	3
mushroom	2	8	2.639	3
nurse	1	5	3.912	2
snake	1	5	3.178	3
spoon	1	5	2.773	1
toilet	2	6	3.367	3

Picture Name	Syllable Length	Character length	Frequency	Age of Acquisition
train	1	5	4.407	1
whistle	2	7	2.303	3
zebra	2	5	1.099	2
sink	1	4	2.773	1
snail	1	5	1.609	3
spider	2	6	2.079	3
umbrella	3	8	2.708	3
tent	1	4	3.807	3
cactus	2	6	1.386	3
cage	1	4	2.833	3
can	1	3	2.303	2
chimney	2	7	2.398	3
deer	1	4	2.565	1
pillow	2	6	2.996	1
ruler	2	5	2.944	3
map	1	3	3.714	3
salt	1	4	3.638	2
smoke	1	5	3.892	3

Notes. These data were extracted from the International Picture Naming Project online-database (Szekely et al. 2003; Szekely et al. 2004).

ii. The target phonemes used in the four experimental conditions (experiment five)

Pictures Name	Arabic Name	Control	Monolingual Condition	Bilingual Condition	Positive Condition
desert	صحراء	/n/	/sˤ/	/r/	/d/
doll	عروسة	/m/	/ʕ/	/l/	/d/
earring	حلق	/f/	/ħ/	/l/	/r/
hair	شعر	/k/	/ʕ/	/r/	/h/
finger	اصبع	/ʒ/	/sˤ/	/b/	/f/
goat	غنمة	/r/	/ɣ/	/m/	/t/
mushroom	فطر	/w/	/tˤ/	/f/	/m/
hammer	مطرقة	/j/	/tˤ/	/m/	/h/
nurse	ممرضة	/w/	/dˤ/	/r/	/n/
onion	بصل	/t/	/sˤ/	/b/	/n/
plate	صحن	/d/	/sˤ/	/n/	/l/
ring	خاتم	/k/	/x/	/t/	/r/
snake	حية	/m/	/ħ/	/h/	/s/
toilet	حمام	/ʒ/	/ħ/	/m/	/t/
train	قطار	/k/	/tˤ/	/r/	/t/
belt	حزام	/ʒ/	/ħ/	/m/	/b/
bone	عظمة	/ʒ/	/ðˤ/	/m/	/b/
bride	عروسة	/n/	/ʕ/	/r/	/b/
chimney	مدخنة	/k/	/x/	/d/	/m/
cloud	غيمة	/s/	/ɣ/	/m/	/k/
drum	طبلة	/w/	/tˤ/	/b/	/d/
eagle	صقر	/s/	/sˤ/	/r/	/l/

Pictures Name	**Arabic Name**	**Control**	**Monolingual Condition**	**Bilingual Condition**	**Positive Condition**
envelope	ظرف	/t/	/ðˤ/	/f/	/n/
flag	علم	/j/	/ʕ/	/l/	/f/
horse	حصان	/w/	/ħ/	/r/	/h/
mosquito	بعوضة	/n/	/dˤ/	/b/	/m/
scissors	مقص	/d/	/sˤ/	/s/	/s/
smoke	دخان	/h/	/x/	/d/	/s/
spider	عنكبوت	/m/	/ʕ/	/n/	/s/
submarine	غواصة	/ʒ/	/ɣ/	/w/	/s/
bread	خبز	/h/	/x/	/b/	/b/
cactus	صبار	/n/	/sˤ/	/r/	/k/
cage	قفص	/t/	/sˤ/	/f/	/k/
can	علبة	/j/	/ʕ/	/b/	/k/
deer	غزال	/w/	/ɣ/	/l/	/d/
frog	ضفدع	/l/	/dˤ/	/f/	/f/
grapes	عنب	/h/	/ʕ/	/b/	/r/
hat	طاقية	/k/	/tˤ/	/h/	/h/
lettuce	خس	/f/	/x/	/s/	/l/
parrot	ببغاء	/ʒ/	/ɣ/	/n/	/r/
salt	ملح	/f/	/ħ/	/z/	/s/
snail	حلزون	/m/	/ħ/	/m/	/s/
table	طاولة	/ʒ/	/tˤ/	/l/	/t/
tent	خيمة	/k/	/x/	/m/	/t/
umbrella	مظلة	/k/	/ðˤ/	/h/	/b/
bird	عصفور	/l/	/tˤ/	/r/	/b/

Pictures Name	Arabic Name	Control	Monolingual Condition	Bilingual Condition	Positive Condition
diaper	حفاظة	/n/	/ħ/	/f/	/d/
donkey	حمار	/f/	/ħ/	/m/	/d/
duck	بط	/w/	/tˤ/	/b/	/d/
hanger	علاقة	/k/	/ʕ/	/l/	/h/
map	خريطة	/ʒ/	/x/	/r/	/m/
nail	ظفر	/t/	/ðˤ/	/f/	/n/
picture	صورة	/j/	/sˤ/	/w/	/k/
pillow	مخده	/j/	/x/	/d/	/l/
ruler	مسطرة	/w/	/tˤ/	/m/	/r/
shell	صدف	/ʒ/	/sˤ/	/d/	/l/
sink	مغسلة	/d/	/ɣ/	/s/	/s/
spoon	ملعقة	/w/	/ʕ/	/l/	/s/
whistle	صفارة	/d/	/sˤ/	/f/	/w/
zebra	حماروحش	/t/	/ħ/	/r/	/b/

iii. Filler pictures (experiment five)

Picture name	Arabic Name	Filler	Syllable Length	Characters Length	Frequency	Age of Acqusition
tiger	نمر	/sˤ/	2	5	2.565	1
dog	كلب	/x/	1	3	4.754	1
fish	سمكة	/ɣ/	1	4	5.100	1
bed	سرير	/dˤ/	1	3	5.136	1
door	باب	/tˤ/	1	4	5.958	1
butterfly	فراشة	/ħ/	3	9	2.398	1
scarf	شال	/ʕ/	1	5	2.565	2
cow	بقرة	/s/	1	3	3.714	1
girl	بنت	/f/	1	4	6.084	1
horse	حصان	/m/	1	5	4.890	1
pig	خنزير	/h/	1	3	3.784	1
chicken	دجاج	/l/	2	7	3.738	1
monkey	قرد	/x/	2	6	2.944	1
bat	خفاش	/l/	1	3	2.708	2
fork	شوكة	/j/	1	4	2.773	1
carrot	جزر	/ħ/	2	6	2.197	1
car	سيارة	/dˤ/	1	3	5.872	1
fireman	رجل اطفاء	/h/	3	7	1.609	2
wolf	ذئب	/f/	1	4	2.398	2
apple	تفاح	/k/	2	5	3.434	1

Notes. The data (i.e., no of syllable and characters, frequency rating and age of acquisition) were extracted from the International Picture Naming Project online-database (Szekely et al. 2003; Szekely et al. 2004).

Appendix J

Characteristics of English and Arabic Consonants (the star (★) indicates shared phonemes)

			Bilabial	Labio-dental	Inter-dental	Alveo-dental	Alveolar	Palatal	Velar	Uvular	Pharyngeal	Glottal
Stop	Voiced	Emphatic				ض /dˤ/						
		Non-Emphatic	★/b/ ب			★/d/ د		★/ʒ/ ج				
	Unvoiced	Emphatic				/tˤ/ ط						
		Non-Emphatic				★/t/ ت			★/k/ ك			
Fricative	Voiced	Emphatic			/ðˤ/ ظ							
		Non-Emphatic								/ɣ/ غ	/ʕ/ ع	
	Unvoiced	Emphatic				ص /sˤ/						
		Non-Emphatic		★/f/ ف		★/s/ س				/x/ خ	/ħ/ ح	★/h/ هـ
Nasal	Voiced	Non-Emphatic	★/m/ م				★/n/ ن					
Liquid	Voiced	Non-Emphatic					★ ر ل /l/r/					
		Emphatic										
Semivowels	Voiced	Non-Emphatic	★/w/ و					★/j/ ي				

Appendix K

Samples of stimuli pictures used in this study (experiments one, two, three, four and five).

Ring

Monkey

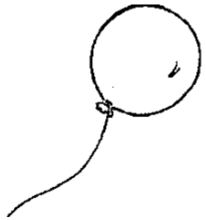
Leaf

Teeth

Map

Strawberry

Spider



Balloon

Notes: All pictures were sampled from the International Picture Naming Project online-database (Szekely et al. 2003; Szekely et al. 2004).