UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Visual Processing of Arabic-English Code-Switching: An Eye-Tracking Analysis

Permalink

https://escholarship.org/uc/item/5rq8t3n9

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

Authors

Shim, Ji Young
Leung, Tommi
Kim, Hyosik

Publication Date

2025

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Visual Processing of Arabic-English Code-Switching: An Eye-Tracking Analysis

Ji Young Shim (jshim@aus.edu)
Department of English, American University of Sharjah, Sharjah, UAE

Tommi Tsz-Cheung Leung (Leung@uaeu.ac.ae)
Department of Cognitive Sciences, United Arab Emirates University, Al-Ain, UAE

Hyosik Kim (hyosik.kim@jj.ac.kr)
Department of English Language and Literature, Jeonju University, Jeonju, Korea

Abstract

This study examines the visual processing of Arabic-English code-switching using an eye-tracking experiment, focusing on determiner-noun switching. Arabic-English bilinguals (L1 Arabic, L2 English) read English-framed sentences that were either monolingual or contained a code-switched noun phrase with an English determiner and an Arabic noun (e.g., *John bought the* كتاب) or an Arabic determiner and an English noun (e.g., *John bought* ال *book*). The Matrix Language Frame (MLF) model (Myers-Scotton, 1993, 2002) predicts that the matrix language (ML) determines the selection of functional elements, requiring determiners to come from English. However, eye-tracking results revealed the opposite pattern: codes-witching was more disfavored if the determiner came from English than if it came from Arabic, contradicting the MLF model's predictions. These findings suggest that grammatical constraints alone cannot fully explain code-switching patterns and that orthographic differences, script directionality, and switching costs play a role in bilingual sentence processing.

Keywords: code-switching; determiner; Matrix Language Frame; Arabic; eye tracking

Introduction

This paper examines the visual processing of code-switching involving noun phrases (NPs) embedded in sentences by analyzing the eye movement patterns of Arabic-English bilinguals. Code-switching, the practice of alternating between two languages within a conversation, is widely recognized as a natural outcome of bilingualism, offering valuable insights into bilingual competence by illustrating how speakers seamlessly navigate between two lexicons (Yow et al., 2018).

Code-switching within NPs has been widely studied, showing that both adults and children frequently switch at grammatical boundaries, particularly between determiners and nouns (Liceras et al., 2008; Mahootian, 2019). Across languages, nouns are the most frequently code-switched grammatical category (Deuchar, 2005; Eichler et al., 2012; Sayahi, 2011). Given this, the grammaticality and processing of NP-internal code-switching have become a primary area of interest in code-switching research and the main focus of the present study (Jake et al., 2002; Joshi, 1985; Lipski, 1978; Parafita Couto & Gullberg, 2019; Poplack, 1980; Timm, 1975). Cross-linguistically, code-switching between a determiner (e.g., *the*) and a noun has been well attested, including Spanish-English, Welsh-English, French-Arabic, and Romance languages such as French, Italian, and Spanish. However, these switches are not random; speakers often exhibit systematic preferences for determiners from one language over the other. This observation has led to proposals about the grammatical and psycholinguistic constraints that shape bilinguals' code-switching competence.

The Matrix Language Framework

Among the many theoretical models, the Matrix Language Frame (MLF) model (Myers-Scotton, 1993) is arguably the most influential in describing code-switching as a phenomenon of language contact (Clyne, 2003; Deuchar, 2006). The MLF model posits that code-switched sentences are structured by a dominant *matrix language* (ML), which determines the overall structure of bilingual sentences. In contrast, the other language, termed the *embedded language* (EL), primarily contributes lexical content. This model assumes an inherent asymmetry between the two languages in code-switched clauses, with the ML exerting grammatical control while the EL provides insertions.

Further refinements of the MLF model introduce additional constraints on code-switching, such as the *System Morpheme Principle* (Jake, 1994; Jake et al., 2002; Myers-Scotton, 1993). According to this principle, the ML determines the allocation of system morphemes (or functional categories), thereby shaping the grammatical structure of code-switched sentences. For example, if there were a code-switch between the determiner and the noun, the determiner is predicted to originate from the ML, while the noun comes from the EL. This prediction has been supported by studies of various language pairs (Deuchar, 2006; Eppler et al., 2017; Herring et al., 2010; Parafita Couto & Gullberg, 2019).

Current Study

Our current study investigated whether the MLF model is applicable in Arabic-English code-switching, a language pair with distinct orthographic systems. While code-switching is common in spoken language, Arabic-English bilinguals are also reported to code-switch in written contexts, particularly in informal settings such as emails, text messages, and social media posts. In these

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* ©2025 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

contexts, bilinguals use both Arabic script and romanized Arabic (often referred to as *Arabizi*), switching not only between languages but also between writing systems (Alsulami, 2019; Warschauer et al., 2006)

Despite its prevalence, research on Arabic-English code-switching remains limited (Altarriba et al., 1996; Alzahrani, 2016; Hamed et al., 2018). One notable study by Hamed et al. (2018) analyzed a corpus of spontaneous conversations by Egyptian Arabic-English bilinguals, revealing that the most common code-switching pattern involved the Arabic definite determiner ال, followed by an English noun. While this finding supports the MLF's predictions, their corpus was predominantly Arabic, leaving open the question of how code-switching operates when English serves as the ML. Additionally, a more controlled approach is needed to examine bilingual speakers' competence in code-switching.

Two key factors make determiner-noun switching particularly relevant to this investigation. First, because Arabic is read from right to left, integrating an Arabic determiner into an English-dominant code-switched NP (e.g., ال *book* 'the book') involves a shift in reading directionality. Conversely, when an English determiner precedes an Arabic noun (e.g., *the* كتاب 'the book'), readers encounter a transition between two different scripts. While both languages follow a determiner-noun order, such script and directionality shifts may create distinct patterns of determiner-noun integration.

Second, English and Arabic differ in how definite determiners are represented orthographically. In English, determiners are independent words, while in Arabic, ال 'the' is a bound morpheme attached to the noun, forming a single orthographic unit. As a result, when ال appears in an English sentence, it remains bound to the Arabic noun, preventing separation (i.e., كتاب ال (book the) is ungrammatical and must be لكتاب (book-the)). These structural differences raise important questions about how bilingual speakers select determiners when code-switching and whether MLF model predictions hold in this context.

Since English determiners and Arabic determiners function differently in terms of morphological attachment, we examine whether determiner-noun switching patterns align with the MLF model. Specifically, we compare:

(1) An English determiner followed by an Arabic noun (e.g., *the* كتاب).
(2) An Arabic determiner followed by an English noun (e.g., ال *book*).

By analyzing these two types of determiner-noun switches, our study investigates whether bilingual speakers systematically follow the MLF model's prediction that the ML (English in this case) dictates determiner selection or whether other structural factors influence code-switching patterns.

Predictions and Hypotheses

The MLF model posits that the ML dictates the structure of code-switched sentences, including the selection of function words (or system morphemes) such as determiners. Applying this model to Arabic-English code-switching leads to the following prediction:

(3) When English is the ML, the determiner should also be English in Arabic-English code-switching.

This prediction assumes that bilinguals adhere to the structural constraints imposed by the ML, even when processing a language pair with distinct orthographic properties. However, given the challenges of reading directionality and morphological attachment, deviations from this expectation may emerge in eye-tracking data.

Although the MLF model is not a processing framework, if its structural predictions extend to online processing, we further hypothesize that:

(4) Eye movement measures will significantly differ between English and Arabic determiners in code-switched noun phrases.

This hypothesis assumes that determiners are processed differently depending on their language, and such differences will be reflected in measures such as first fixation duration, gaze duration, and regressions, metrics that offer insight into how bilingual readers allocate cognitive resources during NP processing. Since code-switching occurs at the NP level, we also predict:

(5) Eye movement measures will differ not only at the determiner but also at the noun.

Differences between switching from English to Arabic (e.g., *the* كتاب) and from Arabic to English (e.g., ال *book*) may reflect how bilingual speakers manage determiner selection while adhering to language-specific morphological constraints.

Experiment

Participants

Initially, 76 Arabic-English bilingual speakers participated in the study; However, four participants were excluded in the final analysis due to either not completing the experiment or not meeting the selection criteria based on the language history questionnaire such as having an age of acquisition (AOA) for either language later than age 7. The final sample consisted of 34 males and 38 females (μ = 19.95 years, σ = 1.71), all recruited from the American University of Sharjah in the United Arab Emirates.The mean AOA was 1 year for Arabic (σ = 2) and 3.3 years for English (σ = 2.3). A paired-samples *t*-test showed a significant difference between the two AoAs, $t(71) = 6.9$, $p < .001$, indicating that participants acquired Arabic

significantly earlier than English. All participants reported engaging in code-switching between the two languages, with all but three doing so habitually; the remaining three indicated occasional code-switching. All participants had normal or corrected-to-normal vision. They received a compensation of USD 16 for their participation. The study received ethics approval from the university's International Research Board, and all participants provided written informed consent prior to the experiment.

Stimuli

A total of 100 English sentences were created, each containing a target NP intended for code-switching, inserted in the middle of the sentence. Each NP began with the definite determiner "the" followed by a noun. All 100 English nouns were high-frequency words, ranked in the top 10% according to Davies and Gardner (Davies & Gardner, 2010). Each target noun was then translated into Arabic, which was also checked for its frequency using Buckwalter and Parkinson (Buckwalter & Parkinson, 2014). Each sentence frame was used across all three code-switching conditions: (1) no code-switching, (2) English *the* + Arabic noun, and (3) Arabic ال + English noun (25 items per condition). An additional 25 sentences featured both Arabic determiners and Arabic nouns, with no switch between the determiner and the noun; these were excluded from the final analysis. no code-switching (25 sentences), code-switching with All stimuli were pseudorandomized using a Latin Square design, with participants viewing each sentence frame in only one of the three experimental conditions. The conditions in (6) describe the three code-switching conditions using the same sentence frame (ENG: English, ARB: Arabic):

(6) a. ENG 'the' + ENG noun (cond: ENG_ENG)
e.g., We sadly missed the train for our winter trip.

b. ENG 'the' + ARB noun (cond: ENG_ARB)
e.g., We sadly missed the قطار for our winter trip.

c. ARB 'ال' + ENG noun (cond: ARB_ENG)
e.g., We sadly missed ال train for our winter trip.

Apparatus

Eye movements of the participants were recorded while reading using the Eyelink Portable Duo (SR Research, Ltd.) with a sampling rate of 1,000 Hz. To minimize head movement, participants used a chinrest. Although participants read with both eyes, only data from the right eye were recorded. Stimuli were presented in black, 20-point Times New Roman font on a white background on a Dell 2N5OTVQ Precision 5760 display monitor (1920 x 1200 x 60Hz). Participants were seated at a distance of 45 cm from the eye tracker and 65 cm from the computer screen displaying the sentences. At this distance, each character subtended approximately one degree of visual angle. Each sentence was displayed on a single line.

Procedure

The experiment was conducted in a quiet, isolated room. Instructions were presented on-screen in text and explained verbally by the experimenter before the session began. A default nine-point grid calibration and validation was performed before starting the experiment. Participants completed seven practice trials and were encouraged to ask questions before proceeding.

The experiment was divided into two blocks, with a 5- to 10-minute break between them. Drift correction was applied at the beginning of each block. Each trial began with a left-aligned fixation asterisk positioned at the onset of the first letter of the sentence. The asterisk disappeared just before the sentence appeared. Sentences were displayed only after participants fixated continuously on the asterisk for at least 500 ms.

Each experiment lasted approximately 40 minutes. Participants indicated the end of each sentence reading by pressing the 'spacebar' on a standard keyboard. After reading each sentence, the participant was asked to rate the sentence in acceptability on a four-point Likert scale worded as *Very Bad, Bad, Good,* and *Very Good*.

Results

Acceptability

To investigate whether the MLF model predicts the language choice of the determiner in Arabic-English code-switching, we analyzed the experimental conditions as follows. To assess the overall impact of code-switching at the NP level, condition ENG_ENG (6a) was compared to conditions ENG_ARB (6b) and ARB_ENG (6c), evaluating whether the presence of a code-switched NP influences acceptability judgments**.**

Next, to evaluate whether the ML determines determiner selection, condition ENG_ARB was compared to condition ARB_ENG. This contrast specifically tests whether participants prefer determiners from the ML, in line with the MLF model's predictions. The analysis was conducted using Helmert contrast coding with the following weights: (0.63, -0.33, -0.33, 0, 0.5, -0.5). Finally, the response levels were categorized into two groups: "(very) good" and "(very) bad", allowing for a clearer evaluation of participants' preferences regarding determiner-noun switching.

A generalized linear mixed-effects model was fitted using the *glmer* function in R (Bates et al., 2015) with maximal convergence settings. The results, presented graphically in Figure 1, revealed that the condition ENG_ENG was rated significantly more acceptable than the conditions ENG_ARB and ARB_ENG ($\beta = 2.45$, $SE = 1.35$, $z = 21.09$, $p < 0.001$), indicating that code-switching generally lowers acceptability.

Additionally, the condition ARB_ENG received higher acceptability ratings than condition ENG_ARB ($\beta = -0.32$, $SE = 0.06$, $z = -3.91$, $p < 0.001$), suggesting that code-switching is more acceptable when the determiner comes from Arabic (EL) than when the determiner comes

from English (ML), contrary to the predictions of the MLF model (see Table 1).

Table 1: Model Comparisons between the condition ENG_ENG and the conditions ENG_ARB and ARB_ENG, as well as condition ENG_ARB vs. condition ARB_ENG, using Helmert contrast coding.

	Acceptability Result			
Predictors	*Odds Ratios*	*std. Error*	*CI*	*p*
(Intercept)	5.72	0.96	4.12 – 7.94	**<0.001**
ENG_ENG	11.64	1.35	9.27 – 14.62	**<0.001**
ENG_ARB	0.72	0.06	0.61 – 0.85	**<0.001**

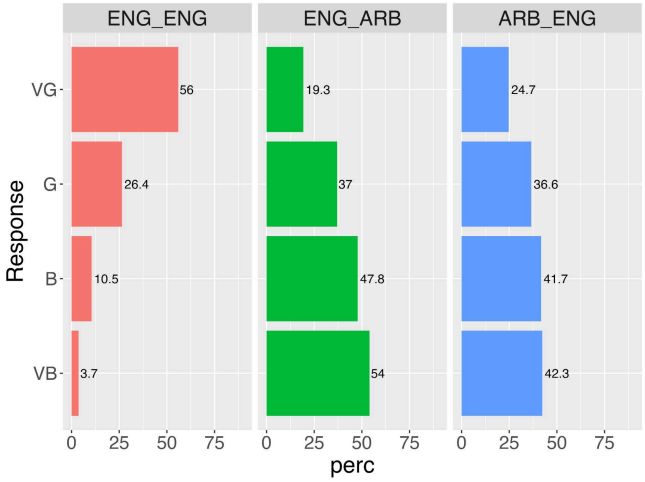


Figure 1: Percentage of responses categorized as: VG(Very Good), G(Good), B(Bad), and VB(Very Bad)

Eye-Tracking Reading Time Results

For the reading time (RT) analysis, the critical regions (target NPs) were divided into two components: determiners and nouns. At the determiner level, the conditions were grouped based on the language of the determiner. Conditions ENG_ENG and ENG_ARB contained an English determiner, whereas condition ARB_ENG contained an Arabic determiner. Similarly, at the noun level, the conditions were grouped based on the language of the noun. Conditions ENG_ENG and ARB_ENG involved an English noun, whereas condition ENG_ARB involved an Arabic noun. Each region was analyzed using Helmert contrast coding to systematically compare the conditions.

Prior to data analysis, single fixations shorter than 100 ms and longer than 1000 ms were excluded. Five reading time measures were reported: *First Fixation* (FF), *First Pass* (FP), *Go-Past* (GP), *Re-Reading* (RR), and *Total Reading* (TR). FF refers to the duration of the first fixation on the target region while FP refers to the sum of all fixations before exiting the target region for the first time. GP is defined as the sum of all fixations from the moment the target region is gazed at until the gaze exits. RR measures the total duration of fixations on the target region following FP, including regressions. TR includes all fixations on the area of interest, regardless of regression. FF and FP are classified as early-stage measures, primarily reflecting lexical access difficulty. In contrast, GP, RR, and TR are considered late-stage measures, capturing overall processing complexity, encompassing both initial processing difficulty and regression/reanalysis.

For the data analysis, a linear mixed-effects model was fitted for each region using the *lme4* function in R with maximal convergence settings. The results are graphically represented in Figures 2 and 3.

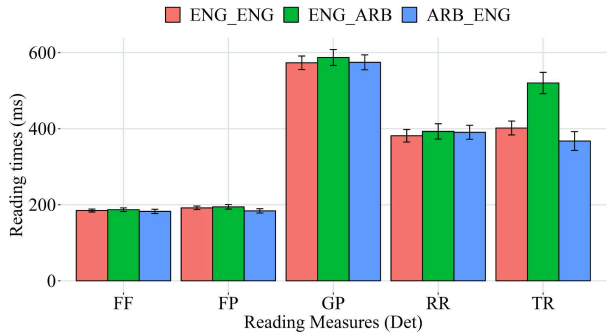


Figure 2: Five eye-tracking measures at the determiner.

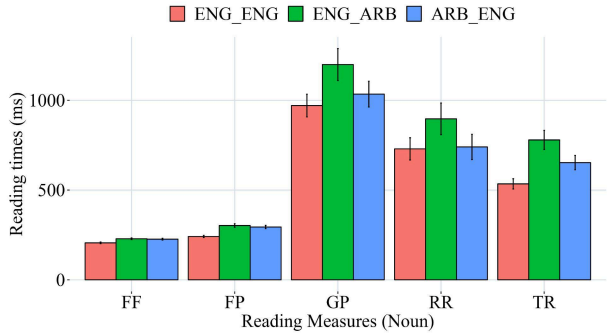


Figure 3: Five eye-tracking measures at the noun.

At the determiner level (Table 2), no significant effect of code-switching was found, except for TR, as shown in Table 2. This finding suggests that code-switching at the determiner level does not significantly increase processing costs.

At the noun level (Table 3), fixation durations were significantly longer across all five reading measures when the noun was Arabic (condition ENG_ARB), compared to when it was English (conditions ENG_ENG and ARB_ENG). This result indicates that code-switching at the lexical category level impacts processing costs, with an Arabic noun being less expected when preceded by an English determiner, particularly when the ML is English.

Additionally, fixation durations increased significantly when the preceding determiner was Arabic (condition ARB_ENG) compared to when it was English (condition

ENG_ENG), as reflected in early-stage measures (FF and FP) as well as TR. This difference may be attributed to the fact that condition ARB_ENG involves multiple code-switchings, e.g., from English (ML) to Arabic at the determiner level, followed by English at the noun level, whereas condition ENG_ENG does not.

Table 2: Model comparisons (determiners)

Model Comparison (HELMERT coding)

	Dependent variable: Reading Measures (Det)				
	First Fixation	First Pass	Go-Past	Re-Reading	Total Reading
	(1)	(2)	(3)	(4)	(5)
ARB_ENG	-4.679	-7.962	-18.911	-11.866	-82.890***
	(5.022)	(5.707)	(28.963)	(28.287)	(27.537)
ENG_ARB	0.933	2.732	23.712	23.117	104.429***
	(5.005)	(5.678)	(28.999)	(28.345)	(27.377)
Constant	185.897***	191.342***	588.935***	399.323***	440.803***
	(3.023)	(3.726)	(15.663)	(14.601)	(18.872)
Observations	716	716	716	716	716
Note:					*p<0.1; **p<0.05; ***p<0.01

Table 3: Model comparisons (nouns)

Model Comparison (HELMERT coding)

	Dependent variable: Reading Measures (Noun)				
	First Fixation	First Pass	Go-Past	Re-Reading	Total Reading
	(1)	(2)	(3)	(4)	(5)
ENG_ARB	11.912***	35.003***	189.750***	154.689***	178.756***
	(2.851)	(4.686)	(28.879)	(28.722)	(16.521)
ARB_ENG	19.067***	51.039***	54.316	3.206	109.676***
	(3.292)	(5.418)	(33.368)	(33.185)	(19.098)
Constant	220.771***	280.139***	1,065.868***	785.695***	653.756***
	(4.202)	(7.034)	(71.734)	(70.385)	(38.943)
Observations	4,748	4,748	4,748	4,748	4,748
Note:					*p<0.1; **p<0.05; ***p<0.01

Discussion

This study investigated whether the Matrix Language Frame (MLF) model accurately predicts determiner selection in Arabic-English code-switching and how bilinguals process code-switched NPs in real time. By analyzing acceptability judgments and eye-tracking reading time measures, we examined whether the matrix language (ML), English in this study, determines the grammatical structure of code-switched NPs, particularly at the determiner level, as predicted by the MLF model.

The results from both acceptability judgments and eye-tracking suggested that determiner selection in Arabic-English code-switching did not fully align with the MLF model's prediction that the ML dictated the grammatical structure of functional elements. Instead, bilinguals exhibited preferences that deviated from cross-linguistically attested patterns, suggesting that additional linguistic and cognitive factors influenced determiner-noun integration in mixed-language contexts.

The acceptability judgment results clearly showed that code-switched sentences generally lowered acceptability. However, condition ARB_ENG, in which the determiner came from Arabic, the embedded language (EL), was rated more acceptable than condition ENG_ARB, where the determiner came from English, the ML. This pattern contradicted the MLF model's expectation that the ML should determine the language of functional elements such as determiners. Instead, Arabic-English bilinguals preferred structures in which the determiner came from their first language (L1) Arabic, which functioned as the EL in this case, rather than the ML English. This suggests that determiner selection in Arabic-English code-switching did not strictly adhere to the hierarchical constraints proposed by the MLF model.

At first glance, the occurrence of the Arabic determiner in isolation is anomalous, as it is orthographically a prefix that should always attach to its host, which is not the case in code-switching. Here we observe a divergent observation between the determiner used in monolingual Arabic and Arabic-English code-switching: in monolingual Arabic, ال /əl/ always functions as a prefix and its pronunciation is subject to further allophonic variation across all Arabic dialects (Leung et al., 2020). On the contrary, in Arabic-English code-switching, the standalone ال is considered acceptable and it is always pronounced as [əl] regardless of the phonetics of the following noun. This suggests that the theoretical status of آل may differ in code-switching, i.e., bilingual speakers may process it as an independent morpheme rather than bound morpheme.

The eye-tracking results further reinforced this pattern. Arabic nouns preceded by an English determiner (e.g., *the* كتاب) elicited significantly longer fixation durations than English nouns preceded by an Arabic determiner (e.g., ال *book*). This asymmetry suggests that bilinguals process code-switching at the noun level differently from the determiner level. One possible explanation is that encountering an Arabic noun in an otherwise English-dominant sentence increases processing difficulty due to script and orthographic differences, requiring additional time for lexical access and integration. This effect may be heightened because the noun carries semantic content and is central to sentence meaning

In contrast, when the code-switched element is limited to the Arabic determiner ال (as in condition 6c), the disruption is reduced, likely because the following noun remains in the ML (English). Grammatically, the combination of ال + a bare English noun is semantically interpretable (i.e., it forms a definite noun phrase). However,, the reverse combination of *the* + a bare Arabic noun can be semantically ambiguous, as Arabic bare nouns are typically indefinite by default, creating a mismatch in definiteness marking Moreover, since ال is orthographically short, its following noun is likely to be parafoveally processed. Bilingual readers may extract basic orthographic cues such as the language of the upcoming word (in this case an English noun), enabling

them to skip the determiner more easily and facilitating a smoother transition across the language boundary.

These findings can be further contextualized within the broader literature on switching costs in bilingual language processing. Some studies suggested that code-switching did not necessarily incur a processing cost, particularly in communities where it was a frequent and natural part of communication (Adamou & Shen, 2019). Other research indicated that switching costs might depend on the direction of the switch, with some studies reporting a higher cognitive load when switching from the dominant or first language (L1) to the weaker or second language (L2) (Bultena et al., 2015), while others found greater costs when switching from L2 to L1 (Faroqi-Shah & Wereley, 2022). The mixed evidence suggests that factors such as proficiency, exposure, and structural properties of the languages play a role in determining the magnitude of switching costs. Our findings contribute to this discussion by showing that increased fixation durations for Arabic nouns following an English determiner suggest a processing cost associated with switching from English (L2) to Arabic (L1). However, the smoother transition observed when an Arabic determiner preceded an English noun indicates that switching from Arabic (L1) to English (L2) may be less disruptive, possibly due to the predictability of English nouns in an otherwise English-dominant sentence.

Further support for our findings came from a recent study on spoken Arabic-English code-switching (Shim & Thaier, 2024), which found that Arabic-English bilingual speakers predominantly preferred the Arabic determiner over the English determiner, regardless of the ML, when code-switched sentences were presented audibly. This aligns with our results in the written modality and reinforces the idea that determiner selection in Arabic-English bilinguals is not strictly governed by the ML, but influenced by additional linguistic and cognitive factors across both spoken and written contexts. The convergence of findings from both the reading and listening modalities suggests that determiner preference may be a robust pattern in Arabic-English code-switching, rather than an artifact of task-specific constraints.

Taken together, our findings challenge the MLF model's assumption that the ML determines the selection of functional elements in code-switched structures. Instead of adhering to some universal constraints, bilinguals may develop language-specific processing strategies that do not necessarily conform to models based on cross-linguistic generalizations.

Future research should investigate whether these findings extend to additional contexts, including spontaneous speech, where factors such as prosodic cues, speaker intent, and discourse context may interact with determiner selection in Arabic-English code-switching. As the reviewers suggested, our hypothesis should also be verified in oral codeswitching, whether it is spontaneous or by elicitation. The use of written code-switching as the object of inquiry has been questioned by reviewers. Callahan (Callahan, 2002) was likely the first to attempt an analysis of written code-switching and to test whether it also supports the MLF model. Her aim was to demonstrate, contrary to critics, that written code-switching is not inauthentic. Moreover, she argued that the MLF model of code-switching can be generalized to written contexts as well. Our study builds on this perspective, and we anticipate that the core findings will be replicable in oral Arabic-English code-switching.

Additionally, individual differences in code-switching such as language dominance may provide further insights into how bilinguals manage grammatical and lexical integration across languages.Another important consideration, as one reviewer noted, is that written code-switching between Arabic and English involves not only a change in linguistic code but also a shift in script, from Arabic (an abjad) and the Roman alphabet. This type of 'script switching' may impose additional cognitive costs independent of linguistic code-switching (Li et al., 2018). Overall, our findings highlight the complexity of bilingual sentence processing and underscores the need for a theoretical model that integrates grammatical constraints, cognitive mechanisms, and language-specific properties, in describing a full spectrum of code-switching patterns.

Acknowledgments

This research was supported by the American University of Sharjah through grant number FRG20-M-S13, awarded to the first author.

References

Adamou, E., & Shen, R. X. (2019). There are no language switching costs when codeswitching is frequent. *International Journal of Bilingualism*, *23*(1), 53–70. https://doi.org/10.1177/1367006917729546

Alsulami, A. (2019). A sociolinguistic analysis of the use of Arabizi in social media among Saudi Arabians. *International Journal of English Linguistics*, *9*(6), 257. https://doi.org/10.5539/ijel.v9n6p257

Altarriba, J., Kroll, J. F., Sholl, A., & Rayner, K. (1996). The influence of lexical and conceptual constraints on reading mixed-language sentences: Evidence from eye fixations and naming times. *Memory & Cognition*, *24*(4), 477–492. https://doi.org/10.3758/BF03200936

Alzahrani, S. (2016). Cross-language semantic similarity of Arabic-English short phrases and sentences. *Journal of Computer Science*, *12*(1), 1–18. https://doi.org/10.3844/jcssp.2016.1.18

Bates, D., Maechler, M., Bolker, B. M., & Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Buckwalter, T., & Parkinson, D. (2014). *A frequency dictionary of Arabic*. Routledge. https://doi.org/10.4324/9780203883280

Bultena, S., Dijkstra, T., & Hell, J. G. V. (2015). Language

switch costs in sentence comprehension depend on language dominance: Evidence from self-paced reading. *Bilingualism: Language and Cognition*, *18*(3), 453–469. https://doi.org/10.1017/S1366728914000145

Callahan, L. (2002). The Matrix Language Frame model and Spanish/English codeswitching in fiction. *Language & Communication*, *22*(1), 1–16. https://doi.org/10.1016/S0271-5309(01)00018-0

Clyne, M. (2003). *Dynamics of language contact*. Cambridge University Press.

Davies, M., & Gardner, D. (2010). *A frequency dictionary of contemporary American English: Word sketches, collocates and thematic lists*. Taylor & Francis Group.

Deuchar, M. (2005). Congruence and code-switching in Welsh. *Bilingualism: Language and Cognition*, *8*(3), 255–269.

Deuchar, M. (2006). Welsh-English code-switching and the Matrix Language Frame model. *Lingua*, *116*(11), 1986–2011.

Eichler, N., Hager, M., & Müller, N. (2012). Code-switching within determiner phrases in bilingual children: French, Italian, Spanish and German. *Zeitschrift Für Französische Sprache Und Literatur*, *122*(3), 227–258.

Eppler, E. D., Luescher, A., & Deuchar, M. (2017). Evaluating the predictions of three syntactic frameworks for mixed determiner–noun constructions. *Corpus Linguistics and Linguistic Theory*, *13*(1), 27–63.

Faroqi-Shah, Y., & Wereley, S. (2022). Investigation of code-switching cost in conversation and self-paced reading tasks. *International Journal of Bilingualism*, *26*(3), 308–333. https://doi.org/10.1177/13670069211056438

Hamed, I., Elmahdy, M. S., & Abdennadher, S. (2018). Collection and analysis of code-switch Egyptian Arabic-English speech corpus. In N. C. et al (Ed.), *International Conference on Language Resources and Evaluation* (pp. 3805–3809). European Language Resources Association (ELRA).

Herring, J. R., Deuchar, M., Couto, M. C. P., & Quintanilla, M. M. (2010). "I saw the madre": Evaluating predictions about code-switched determiner-noun sequences using Spanish–English and Welsh–English data. *International Journal of Bilingual Education and Bilingualism*, *13*(5), 553–573.

Jake, J. L., Myers-Scotton, C., & Gross, S. (2002). Making a minimalist approach to code-switching work: Adding the matrix language. *Bilingualism*, *5*(1), 69–91.

Joshi, A. (1985). Processing of sentences with intrasentential code-switching. In D. Dowty, L. Karttunen & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational and theoretical perspectives* (pp. 190–204). Cambridge University Press.

Leung, T. T.-C., Ntelitheos, D., & Al Kaabi, M. (2020). *Emirati Arabic: A comprehensive grammar* (1st ed.). Routledge. https://doi.org/10.4324/9780429273162

Li, R., Zhang, Z., Ni, C., Xiao, W., Wei, J., & Dai, H. (2018). Examining the functional category in Chinese–English code-switching: Evidence from the eye-movements. *Journal of Psycholinguistic Research*, *47*(1), 1–28. https://doi.org/10.1007/s10936-017-9513-z

Liceras, J. M., Fernández Fuertes, R., Perales, S., Pérez-Tattam, R., & Spradlin, K. T. (2008). Gender and gender agreement in bilingual native and non-native grammars: A view from child and adult functional–lexical mixings. *Lingua*, *118*(6), 827–851.

Lipski, J. (1978). Code-switching and the problem of bilingual competence. In M. Paradis (Ed.), *Aspects of bilingualism* (pp. 250–264). Hornbeam Press.

Mahootian, S. (2019). *Bilingualism*. Taylor & Francis Group.

Parafita Couto, M. C., & Gullberg, M. (2019). Code-switching within the noun phrase: Evidence from three corpora. *International Journal of Bilingualism*, *23*(2), 695–714.

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en Español: Toward a typology of code-switching. *Linguistics*, *18*, 581–618.

Sayahi, L. (2011). Code-switching and language change in Tunisia. *International Journal of the Sociology of Language*, *2011*(211), 113–133.

Shim, J. Y., & Thaier, L. B. (2024, November 6). *AI vs. Human: A preliminary investigation into Arabic-English code-switching*. Paper presented at The Sharjah International Conference on AI & Linguistics, American University of Sharjah, UAE.

Timm, L. A. (1975). Spanish–English code-switching: El porqué y how-not-to. *Romance Philology*, *28*(4), 473–482.

Warschauer, M., Said, G. R. E., & Zohry, A. G. (2006). Language choice online: Globalization and identity in egypt. *Journal of Computer-Mediated Communication*, *7*(4), 0–0. https://doi.org/10.1111/j.1083-6101.2002.tb00157.x

Yow, W. Q., Tan, J. S. H., & Flynn, S. (2018). Code-switching as a marker of linguistic competence in bilingual children. *Bilingualism: Language and Cognition*, *21*(5), 1075–1090.