

Lexical Borrowing in Arabic and the Role of Orthography

Sami Hamdi

University of New Mexico
Jazan University
English Language Center

Abstract

Explanations of how loanwords are adapted and integrated in the recipient language center mainly on phonological and morphological accounts. There are various factors that mediate the adaptation process and worth further investigation. Upon the review of literature, it was observed that some loanwords have multiple written forms in Arabic. A study was conducted on 13 Arabs bilingual in English to examine the roles of orthography on loanwords adaptation by bilinguals. The results revealed a salient role for orthography in adapting single and compound loans affecting certain phonological and morphological aspects of L1. Arab Bilinguals were sensitive to the orthography of loans when they adapt them to their L1 basing their spelling choice on standard Arabic. The findings suggest that the adaptation process is phonetic in nature and influenced by orthography resulting in multiple written loan-forms.

Keywords: Loanwords, loanword adaptation, lexical borrowing, Arabic, orthography

1. Introduction

Loanwords are one of the most common phenomena in language contact and almost every language exhibits one or more forms of borrowing. In lexical borrowing, words are transferred from one language and integrated into another language. Haugen (1950) introduced a notable taxonomy to distinguish between different borrowed items:

- *Loanwords* involve copying both the form and the meaning
- *Loan blends* are those borrowed words where a copied part exists along with a native part
- *Loan shifts* show copying only of the meaning and include both *loan translation* and semantic borrowing.

Thomason & Kaufman (1988: 21) confine borrowing to “the incorporation of foreign elements into the speakers’ native language”. In this sense, borrowing is argued to be *interference* and non-native speakers borrowing is not included. Loanword adaptation involves the phonological and morphological transformation of foreign items to fit the grammatical system of the recipient language. However, according to Haspelmath (2009: 43), “The precise ways in which the adaptation process happens are often complex and a matter of ongoing debate”. The extent to which loanwords conform to the recipient language differs from one language to another. In other words, loanwords might adhere to the recipient language system’s phonology and morphology in some respects, but they might conflict with other patterns. Several theories (see in 2.) have been proposed to account for how loanwords are adapted and incorporated in the recipient languages.

The debate on loanword adaptation focuses on whether the process by which loans are transformed to another language is phonetic or phonological. There are, of course, differing sociolinguistic and psycholinguistic factors that come into play during integrating loans in the recipient language, including the motivation for borrowing and the ideologies about the native and other languages (see Sections 2.4 and 2.5 for a review of this literature). It comes as no surprise to see the studies on loanwords involving bilinguals or second language learners. This is primarily due to the key role bilinguals and second language learners play in both introducing and adapting loans in the recipient language. Thus, insights from the literature on bilingualism can inform explanations of loanword adaptation, both for how speakers incorporate foreign words grammatically and also how they view and deal with them in their native languages. Extra-linguistic factors such as orthography are also important, though they are much less often investigated than other factors.

Although loanword adaptation is mainly a grammatical phenomenon, these extra-linguistic factors intervene in their incorporation into the recipient language and require further. A careful review of literature shows that the role of orthography is often disregarded and, when it is mentioned, it is marginalized. This paper, reviews research on loanwords mainly in Arabic and addresses the role of orthography on loanword adaptation in order to fill this gap in the literature. A set of loanwords that have multiple written forms in Arabic, were presented to a group of native speakers of Arabic who are also fluent in English. These speakers were asked to rate and comment on the different forms. Results demonstrate that multiplicity in orthographic forms is not a matter of free variation but is instead shaped by factors such as speakers' awareness and frequency. This paper is organized as follows: In Section 2, explanations of loanword adaptation and the sociolinguistic factors behind lexical borrowing are presented. In Section 3, an overview of the literature on lexical borrowing on Arabic is introduced in terms of phonetic and phonological, morphological, and semantic aspects of adaptation as well as sociolinguistics. Section 4 describes orthography role in previous studies and the motivation for examining it in Arabic. The study questions, design, and findings are reported in 5 and the conclusion in 6.

2. Loanword adaptation: explanations and sociolinguistic implications

2.1 Adaptation theories

Phonetic and phonological explanations have been seen as the major accounts of loanword adaptation. There are three positions that describe the phonetic and phonological adaptation of loanwords. Within the phonetic view, input is assumed to consist of acoustic signals which are the surface forms of loans mapped onto native phonemes based on similarity. Advocates of the phonetic theory (Silverman, 1992; Peperkamp & Dupoux, 2003; Peperkamp, 2004; Kabak and Idsardi, 2007) argue that speakers have no access to the donor language phonological system and the adaptation process is thus solely phonetic in nature. For example, Silverman (1992) points out that Cantonese bilinguals are unable to access phonological structures of the loanwords.

The phonological view of adaptation, on the other hand, argues that speakers perceive foreign sounds within the framework of their native phonological system and, consequently, transform these foreign elements into their nearest native correspondents. Advocates of this position (Paradis & LaCharite 1997; Jacobs and Gussenhoven, 2000; LaCharite & Paradis 2005; Itô & Mester 1995) downplay the effect of phonetic encoding, claiming that bilinguals' access to both the donor and recipient language phonological systems suggest that adaptation is a phonological process rather than simply a phonetic one. Under this view, an L2 segment is replaced by the closest phonological, but not phonetic, unit in the L1. An example to show this view is English high lax vowels /ɪ, ʊ/ in Mexican Spanish which are adapted based on phonological rather than phonetic similarity. Chang (2008) and others (Broselow, 2004; Kang, 2003; Kenstowicz, 2001/2004; Shinohara, 1997; Steriade, 2001; Yip, 1993) propose an intermediate account that includes both the phonetic and phonological levels of adaptation. For instance, Chang (2008) cites examples from English loans in Burmese and points out that bilinguals' use of English loans in Burmese is characterized by phonological and phonetic scansion of the L2 input. These explanations show how complex loanword adaptation is, especially when theorizing is restricted to particular aspects. Extra-linguistic factors such as mispronunciation and mishearing remain challenges for these theories.

Campbell (2004) accounts for the disorderly substitutions of sounds in borrowed words with two explanations. First, he points out that loan adaptation differs based on the time of borrowing: older loans incorporate sounds before language contact intensifies and newer segments appear in more recent loans. He refers to Sayula Popoluca *turu* 'bull' which was borrowed from Spanish *toro* and notes that *r* is usually replaced by *n* in prior loans. Campbell's second explanation is associated with the orthography, where the pronunciation is often based on spelling conventions. Poplack, Sankoff, and Miller (1988) agree that the length of time loans exist in the language is a significant factor and they also demonstrate that the frequency of use of a particular item is important in the degree of its phonological adaptation. Loans, they say, more closely match the L1 phonological system when they are more frequently used by a large number of speakers over time.

2.2 Sociolinguistic factors for lexical borrowing

The sociolinguistics of lexical borrowing include a plenty of factors that account for different linguistic and sociocultural contexts.

There are mainly two reasons for lexical borrowing: *need*, what can be called cultural borrowing, in which a language does not have a word for a given concept (e.g.; Arabic *coffee* in many languages and Spanish *arrusa* ‘rice’ in Quechua) and *prestige* where speakers use words that have equivalents in their native language which may be motivated by socioeconomic status factors (e.g.; social class, education level, income). Another, but less common, reason is borrowing due to taboo where some words have derogatory senses. The Haruai case in Comrie (2010) explained (see 2.3) is an example of this negative association or perception. Other examples are Korean *hostis* (female working in bar) from English ‘hostess’ and French *hâbler* (to brag) from Spanish *hablar* ‘to speak’. The rate of lexical borrowing varies in different contact situation. Sociocultural factors such as power, identity, and ideology play a significant role behind the degree of lexical borrowing. Conquest and immigration are two situations where power of the dominant group is exercised so that less dominant groups acquire linguistic items from the dominant language under pressure. For instance, European invasions to Americas and immigration contribute to the imposition of lexical items in the languages of subordinate groups; i.e. many North American languages. The Vaupés river basin area (in the Upper Rio Negro region of the northwest Amazon), which is linguistically diverse and rich, exhibits one of the striking example of the affiliation between language and identity. Although linguistic exogamy is practiced in this area where inhabitants are required to marry females from other languages, lexical borrowing is relatively limited (Epps, 2013). This lexical avoidance demonstrates not only the maintenance of linguistic distinction but also the view of language and identity as the same. The languages of the Chaco region as described by Campbell and Grondona (2010) resist lexical borrowing despite the intense contact with other languages. This resistance is exercised by ideological practices where individuals who claim to speak only one language and that they understand but not speak the other languages; the so called ‘dual-lingualism’. Obviously, these sociocultural factors are difficult to measure but also their influences can’t be denied.

3. Lexical borrowing in Arabic

Arabic exhibits a rich content of loanwords mainly nominal from different languages such as Turkish, Persian, French and English. When studying lexical borrowing in Arabic, it is important to discuss Modern Standard Arabic (MSA) in relation to other varieties. MSA is the modern form that developed from classical Arabic (CA); the language of religious teaching and early literary work. It is used in writing, news, and formal communication across Arab world. Other varieties of Arabic such as Egyptian Arabic (EA) and Jordanian Arabic (JA) spread as regional dialects with different degrees of intelligibility cross dialectally. Loanwords in MSA and other Arabic varieties demonstrate different adaptation informed by phonotactics and derivational/inflectional patterns due mainly to applying native grammar. Arabic and MSA will be used henceforth interchangeably, but reference will be made to varieties when needed.

3.1 Phonological adaptation of loans in Arabic

Arabic has twenty-eight consonantal phonemes and six vowels. Although nativizing foreign elements might be predictable to some extent, it may violate the borrowing language system as the situation in many cases in Arabic. The degree of phonological adaptation varies in MSA and the various Arabic dialects effecting differences in pronunciations. For example, *doctor* might be pronounced /dʌktɔːr/, /daktuːr/, or /daktɔr/. However, these dialectal differences in pronunciation exhibit similar patterns in general. There are three major types of phonological adaptation of loanwords in Arabic: *replacing same sounds by other sounds of the same natural class*, *non-Arabic sounds adaptation*, and *vowel epenthesis*.

Foreign sounds may change into their Arabic counterparts which are not necessarily the native correspondents. Table I illustrates some sounds that are replaced by other sounds from the same natural class.

Table I: foreign sounds and their Arabic counterparts

Foreign sound	Arabic counterpart	Example loanword	Source language	Loan-form
/t/	/tʕ/	tulumba ‘pump’	Latin	/tʕulumba/
/d/	/dʕ/	moda ‘fashion’	Italian	/mudʕə/
/ʌ/	/a/	bus	English	/basʕ/
/o/	/u:/	mall	English	/mu:l/

Despite anomalies, non-Arabic consonants are often adapted by their nearest equivalents through certain processes such as replacement, addition, and omission. We will present cases of only four foreign consonants /p, tʃ, g, v/. For instance, /p/ >/b, f/ in barkayh parquet and ʔisfanʒ ‘sponge’. The /tʃ/ >/ʃ/ and is retained by some speakers in fayk ‘cheque’. Also, /g/ >/ʒ, ʒ, k/ inʒas ‘gas’, sidʒa:rah ‘cigarette’. In vanilla and vitamin, /v/ >/f/ as fa:nillaand fitami:n.

Vowel insertion is one of the typical processes to break consonant clusters in loanwords. It is a phonotactic mechanism in order for consonant clusters to conform to native phonological patterns. This vowel insertion can be word initially, medially, and finally as in *aspirin* > /ʔisbiri:n/, *cravat* > /karafat/, and *ounce* > /ʔu:nsa/. It is important here to mention again that when vowels occur in Arabic word initially they are accompanied by a glottal stop, yet they can be treated technically as vowels. Other types of phonological changes are also observable such as vowel deletion as in *Christmas* > /kirismis/ and stress shift in *mechanic* > /mikan'i:ki/.

The phonological changes aforementioned show a reasonable degree of regularity though not always the case. For example, the initial [s] in words like *Caesar* > q /qaysʕar/ seems contradicting the norms. In other words, although there is a fair degree of regularity, there is no one to one mapping phonological criteria to predict the exact corresponding phoneme. Yet, some attempts were made to explain this irregularity of some foreign sounds changes in the presence of their Arabic equivalents. According to Al-Qinai's (2000), Arabic sounds often occur in particular combinations which suggests a reasonable degree of harmony. For instance, in *Christmas* > /kirismis/ the initial consonant cluster is broken by [i] and the [a] shifts to [i] based on a principle of vowel harmony with the previous one. Another explanation is related to emphatic sounds which are one of the distinguishing phonological aspects of Arabic where early Arabs tried to maintain in the borrowed items. Al-Qinai refers the exceptional cases to two factors: level of education where individuals try to retain source sounds of loanwords and the influence of dialect due laziness. However, some loanwords cited by Al-Qinai are old and not in use currently, beside the fact that they were standardized by Arab philologists based on standard Arabic which may justify their higher degree of regularity.

3.2. Morphological adaptation of loanword in Arabic

The morphological adaptation of loanwords addresses mostly word derivation, number assignment, and gender assignment which are the most important morphological aspects in the integration of foreign words into Arabic. Common word derivation processes are employed to incorporate foreign words in Arabic such as clipping, compounding, remodeling, derivation and inflection as in the following examples from Al-Qinai (2002): *baridah dam* (from Persian) ‘mail’ > /bari:d/, *baking powder* > /baykinbawder/, *patrikos* (from Greek) ‘penguin’ > /batri:q/, and (non-analogical) /hawun/ ‘mortar’, and *cable* > kabilat, kibalat, or kawabil. These loan words may or may not undergo analogical remodeling or show inflectional affixes in accordance with Arabic morphological system. There is no strong tendency towards defining criteria of analogical modification in order for loans to undergo and this can be seen from the multiple forms loans may take when Arabized.

Number and gender are core grammatical categories in Arabic. Loanwords are assigned number based on animacy following the native grammatical pattern. The singular is the base form which is unmarked as in *film* and *albu:m* from English. In MSA, the dual form has the suffix /-a:ni/ and /-ayni/ based on the case. The plural has mainly two methods: the *sound plural* which applies the suffix /-u:n/ and /-i:n/ depending on the case to code masculine human referents as in musi:qiy-u:n/i:n for *musicians* while the suffix /-a:t/ and its variant /-ha:t/ mark the feminine in kamira:t and kazuha:t for ‘cameras’ and ‘casinos’. The other method is called *broken plural* which is basically the ablaut where vowels are inserted within the word to form the plural but do not have a regular predictable pattern as in afla:m and dakatira for ‘films’ and ‘doctors’. As to gender, Arabic has a masculine and a feminine gender where the assignment of gender is based on the semantic aspect of animacy. If the referent is animate, the unmarked form is the basic which is masculine whereas the feminine is coded by the suffix /-ta:/ and its variants /-a/ and /-ha/ i.e. brufisu:r and brufisu:r-a for male and female *professor*. When the referent is inanimate, gender assignment is believed to be irregular. When it comes to the varieties of Arabic, there are differences in assigning number and gender to foreign words. EA partially follows MSA applying the suffix /-i:n/ for masculine and /-a:t/ for feminine as well as the broken plural inflection that is unpredictable taking several forms. Hafez (1996) noticed that some loans in EA have limited or partial derivation and for gender assignment, loans are inflected in accordance with their original marking as in *cashier* > kaʃi:r(m.) and kaʃi:ra (f.). However, other loans may not inflect for feminine such as ‘mechanic’. She further explains that loans with inanimate referents tend to inflect for the feminine.

She attributes the degree of integration of loans to several factors such as their ability to conform to the grammatical system, whether they agree with their corresponding homonymy, frequency of usage, and speakers' attitude toward foreign words. Table 2 from Hamdi (2017) illustrates variation in number assignment in Arabic varieties.

Table 2: Plural forms of loans cross dialectally

Loanwords	MSA pl.	Broken pl. in varieties
lorry	lu:riyya:t	lawa:ri (EA)
radio	ra:djuha:t	rada:wi (EA)
gallon	dʒalu:na:t	galaneIn (JA) or dʒawali:n in other varieties
villa	filla:t	v/fillal (EA)
blouse	blu:sa:t	bala:jIz (JA)
cigarette	sidʒar /-a:t/-ʔir/-jIr/	saga:jIr (JA)
machine	ma:kina:t	ma:kina:t/ ma:ka: ʔin

In JA, gender assignment is based on the phonetic ending and on the referent's sex. For instance, the final *a* in *camera* and *pizza* is analyzed as a feminine marker following native grammar while /brIntar/ 'printer' is treated as feminine based on their equivalent in Jordanian Arabic. Alsaidat (2011) put it that all inanimate non-singular nouns are inflected for femininity regardless of the gender in their singular forms. Besides applying the native pattern or the nearest equivalent, Hamdi, (2017) draws attention to a semantic factor in that differences in gender assignment cross dialectally is based on the sense of the referent. For example, *cream* is masculine when the referent is *ointment* while feminine when the referent is *whipped cream* in consistence with their nearest native equivalents *marham* and *qiʃda*. Poplack, Pousada, and Sankoff (1982) look at variation on gender assignment as language specific than universal that might be explained at initial assignment of gender.

Most loaned compounds in Hadrami Arabic were contracted into single words and with the second part omitted in some cases as in *kni:cab* 'knee cap' and *self* from 'self-starter' respectively (Bahumaid 2015). Furthermore, number and gender assignment tend to generally follow native patterns except for some irregular cases. In Hadrami Arabic, some loan nouns give rise to verbs forms whereas some loan verbs derive other forms as in 'goal' > *gawwal* 'to score a goal' and 'finish' > *fannaʃ* > *tafni: ʃ*(verbal form) > *finniʃ* (imperative) respectively. Both MSA and the varieties show resistance to foreign inflection. Poplack, Sayahi, Mourad, and Dion (2015) found that French nouns behave following their counterparts in Tunisian Arabic. They associate this behavior with the semantic imperative of expressing plurality and avoiding inflection. The morphological adaptation of loans in Arabic seems to be complex requiring various processing to conform to native grammar. Although it is a universal tendency to see more nominal loans than verbal, yet in Arabic verbs along with their derived forms are the most basic lexical elements from which a considerable amount of lexical entries are generated. Furthermore, Arabic has more inflectional categories. The English word *filter* has a quadrilateral consonant root in Arabic *f-l-t-r* and can have the derived forms: verbal noun *faltarah*, an active participle *filtar*, and usually a passive participle *mufaltar* but is incapable to produce more lexical entries. Nominal loans are treated as solid stems in Arabic which are unanalyzable into roots and patterns.

3.3 Semantics

The literature on the semantics of loanwords in Arabic focuses on semantic domains and semantic changes. Al Athwary (2016) investigated the semantics of English loanwords in MSA media language and found that computer and technology, politics and military, and medicine as the most frequent domains. The findings exhibit common semantic change mechanisms such as extension, restriction, amelioration, pejoration, and metaphorical extension. He points out that semantic change is characterized by restriction more than other mechanisms. He refers the scarcity of semantic shift in Arabic media language to the regular adaptation of loans by institutional efforts to incorporate in MSA. Likewise, Yaseen (2010) studied the frequency and domains of English loanwords Arabic media and found that technology and abstract concepts were the most frequent domains of loans with restriction as the main semantic change. The science and technology was the major semantic field in Palastinian and Jordanian Arabic (Butros', 1963). The semantic change of English and French loanwords in Jordanian Arabic is characterized by widening in Baders' (1990) list.

Egyptian loans from French draw heavily on the semantic domains of furniture, art, and fashion with technology as the main domain of English loans during late eighties and nineties (Ibrahim, 2006). Restriction was the chief semantic change in Hadrami Arabic as Bahumaid (2015) found who also noted that some loans which are semantically restricted can be further associated with specific fields of usage as *back* and *center* in football or narrowing the referent of the loans as in *coffee* which refers to Western-style coffee.

The findings from research on semantics agree with a universal trend where most of loans are nominal. The domain of science and technology reflects a lexical gap in Arabic which explains why most loans fall under that domain. Media has been an essential source of introducing loans and this shows the role of media as a powerful means to reach out public. Semantic change of loanwords in Arabic can be characterized by restriction or narrowing. This might be justified by speakers' tendency to use certain loans with particular senses and exclude unnecessary senses as in *security* only for security guard/officer and *routine* only for a regular procedure/action of doing things.

3.4 The sociolinguistics of loanwords in Arabic

The sociolinguistics of loanwords in Arabic can be approached in terms of code-switching, variation, and sociocultural factors. Code-switching illustrates bilinguals and L2 learners/speakers behavior to alternate between the two languages using single words or phrases. Mustafawi (2002) examined lone English-origin nouns in Arabic contexts. She found that lone items operate according to the recipient language grammar and they are better seen as borrowings than code-switching. In Bahraini Arabic, students apply rich morphological modifications to borrowed words such as definite article and pluralization (Alsadeqi 2010). Applying native grammar to foreign words is normal but does not necessarily mean these foreign words are loans. Code-switching might be better seen at the phrase level than single word level. One possible way to distinguish code-switching from loans is to see the adaptation level. Code-switching does not show adaptation of phonological and morphological features as in well-established borrowing and, thus, such non-conventional words are called nonce-borrowing (Sankoff, Poplack, and Vanniarajan, 1990). Nonetheless, the characterization of loans and code-switching has been a controversial issue in general.

Code-switching has been associated with positive attitude towards the target language. Hussein (1999) addresses university students' attitudes toward code-switching and code mixing with English in their Arabic discourse and describes both negative and positive attitudes. However, he put it that the main reason for code-switching was the absence of readily Arabic equivalents. Prestige and habits are salient factors behind using loanwords. Al Btoush (2014) reports that English loanwords are used by students of scientific faculties more than humanities as well as by females more than males. This partly demonstrates a justified practical application of loans in Arabic especially in science and technology where native equivalents may not be available. Arabic dialects show variation in using and integrating loanwords in their native varieties. For example, *film* and *petrol* are realized /filim/ and /bɑtrɔ:l/ in other dialects. Similarly, *gallon* can have three plural forms in Arabic varieties: /dʒalu:na:t/, /galaneIn/, and /dʒawali: n/. Variation is also in the use and choice of loans that have similar concepts and multiple synonyms. English *fresh* and Turkish *taza* both denote something newly made/produced, also French *as ensure* and English *lift* denote a device for carrying people or an elevator. Different factors have been proposed to explain variation such as speakers' sluggishness, degree of bilingualism, and level of education. The variation is informed by differences in dialects which have a reasonable degree of regularity based on the native phonological system of the dialects and this show two more factors, frequency and context. Frequency is a common factor in the adaptation patterns; in that loans can follow or violate local patterns based on how they initially spread through the lexicon. Variation in adaptation or lexical choice is likely constrained by the need to communicate following the community conventions. Therefore, the different social groups use the native forms or patterns of their varieties. Whenever the context changes, as being in a formal context, standard form will likely be used. Also, when speakers predict that the interlocutors do not have access to the local form, they most likely switch to the standard or a more frequent form from cross-dialectal.

The motivations for code-switching and variation in the use and pattern of loanwords incorporation can be further highlighted by sociocultural aspects such as power, ideology, and identity. Power can be seen in the variation of loanwords that has the same similar concepts. The previous example loans expressing similar concepts are used by different social groups not merely for differentiation, but because they were under different colonial powers. For instance, many loans in Tunisian Arabic are French while many other loans in Egyptian Arabic are English.

Obviously, other Arabic varieties were also in contact with Persian and Turkish and imported a lot of loans due to sociopolitical power of different cultures that dominated Arabs world. On the other hand, ideology is reflected on speakers' choice of borrowed items. By looking at the semantic domains of loanwords, one can see that religious concepts are lacking. Here, speakers' avoidance is ideologized as demonstrated by their rejection to use foreign words that have religious denotations. This avoidance may be due to speakers' strong belief in their religion which is mainly Islam and that such borrowing is not needed and, thus, unacceptable. The avoidance of particular borrowed forms and the enforcement of certain adaptation patterns (that differ from other varieties) are instances symbolizing identity preservation. Thus, speakers' preferences or reservations against some forms are motivated by showing loyalty and respect to their social groups and identity. This is also applies to bilinguals and L2 who code switch among each other as symbol of shared identity, being college students, and where the interlocutors are monolingual, code-switching is minimized if not discouraged.

4. Orthography in loanword adaptation

Part of the transformation process of loanwords in the recipient language is the spelling or written form. Many researchers (see Haugen, 1950; Campbell, 2004; Peperkamp, 2004; Kenstowicz and Suchato, 2006; Paradis and LaCharité, 2008; Dohlus, 2005) suggest a role for orthography during adaptation, yet only few have examined it. Vendelin and Peperkamp (2006) is one of the notable experimental works to examine orthography impact on loanwords. They studied loans adaptation among French listeners and illustrated that the presence of English orthography played a significant role in adapting vowels in the French forms. Detey and Nespoulous (2008) examined the syllabic segmentation of non-words to see the perception of tautosyllabic consonant clusters by Japanese learners of French. The task was presented in three conditions: auditory, visual and synchronous audiovisual. A plausible influence of orthography occurs on the audiovisual and visual conditions effecting more epenthesis than the auditory condition which was interpreted as due to more phonological than phonetic representations. Following Vendelin and Peperkamp (2006), Daland, M., and Kim (2015) investigated the adaptation of English vowels in Korean and found that English orthography guides Korean vowels adaptation especially unstressed ones. The reasons why orthographic effect has not been found in earlier studies or where it was expected to occur are explained by Daland, M., and Kim (2015: 74) "The first is that orthographic effects are pervasive throughout an adaptation system but have subtle and/or variable consequences, so that they can only be detected by large scale statistical studies or carefully targeted experimentation. The second possibility is that orthographic effects are restricted to particular phonological/perceptual contexts, and become evident only when these particular contexts are studied in detail". These studies, though limited, indicate a possible effect of orthography in that the inclusion of orthography can condition how loanwords are adapted.

4.1 The role of orthography in loanword adaptation in Arabic

The goal of this study is to investigate the role of orthography in loan word adaptation in Arabic, a topic that has been neglected in the literature. Although there is a strong correspondence between the sounds and spelling (how words are pronounced and written) in Arabic along with a limited number of vowels, many loans have more than one written forms not only in dialects but in MSA as well. Orthography is hypothesized to play a role in the multiple loan-forms in Arabic. For instance, *chocolate* have *fuklata* and *fuku:lata* as well as the word *computer* have *kombju:tAr* and *kambju:tAr*.

4.2. Aims of the present study

The multiple written forms of loans in Arabic suggest that orthography may play a role in loan adaptation in the recipient language. Although all multiple loan-forms for loans are present in Arabic, one form may be described a more foreign with the other as more native. This study was intended to provide a look at how bilinguals decide on the orthography of loans with multiple written forms. In particular, the study aimed to address the following questions:

1. Do bilinguals prefer a more native or source orthographic form for loans?
2. Are bilinguals more aware to analyzability or phonology during loan compound adaptation?

The rationale behind the study is that the choice of orthographic form may provide us with insights on how grammar operates during phonetic/phonological processing of loans and whether loan compounds are perceived as single unites or sequence of words.

5. Methods

5.1 Sample and procedures

The target population in this study was Arabic speakers who have become bilingual in English later in life. A convenient sample of Arab individuals who are late bilingual in English (and known to the researcher) were selected and invited to participate in a survey by email. A total of 13 Arabs bilingual in English participated in this study, ranging in age between 20-40. Four of the participants were female and eight were males.

5.2 Materials

A questionnaire was designed and validated by a professor in linguistics. The questionnaire was set online through *Google Forms* and sent as a link to the participants by emails. Consent was obtained before each participant began the survey and the survey assessed socio-demographic information including age, gender, level of education, and language background as an optional measure. All of the participants responded and filled out the online survey. A set of ten English loanwords were presented: six of these were single words (*diploma*, *cabin*, *biscuit*, *dolphin*, *catalog*, *sandwich*) and four were compound loans (*icecream*, *hamburger*, *microphone*, *cheesecake*). The choice of these loans was based on having two written forms in standard Arabic for each loan. Loans were checked for multiple written forms and collected from arTenTen12 corpus (Belinkov et al. 2013) through Skitch Engine website. One loan-form is seen to correspond to the source word orthographically more than the second. Participants were given the two written loan-forms for every loanword in Arabic and were asked to choose one better written form. For example, one question reads *What is a better written form of the word 'biscuit' in Arabic?* followed by two written forms in Arabic: بسكويٲ baskawi:t and بسكوت basku:t.

6. Results and Discussion

The results show that Arabic orthographic forms of loans were influenced by English spelling. Three of the loans (*cabin*, *biscuit*, *sandwich*) were chosen by participants as better be written conforming more to English orthography, loan-form 1. The other three loans (*diploma*, *dolphin*, *catalog*) were chosen in loan-form 2 that is less similar to the source loan spelling. Table 3 below illustrates single loan entries with the most selected loan-forms in bold along with the number of responses of the most selected form:

Table 3: single loanwords with multiple written forms¹

Loanword	Loan-form 1	Loan-form 2	# of responses
diploma	دبليوما diblu:ma	دبليوم diblu:m	10
cabin	كبينة kabi:na	كابينة ka:bi:na	9
biscuit	بسكويٲ baskawi:t	بسكوت basku:t	8
dolphin	دلفين dulfi:n	دولفين du:lfi:n	9
catalog	كتلوج kataludʒ	كتالوج kata:ludʒ	11
sandwich	ساندوتش sa:ndwitʃ	سندوش sandawi:f	12

The selection of loan-form 1 is triggered by the orthography of the source loans in English. For *biscuit*, participants preferred the loan-form baskawi:t, which is near the English spelling, to loan-form basku:t. Thus, the influencing factor comes from English spelling which can be clearly seen in the correspondence between the second syllable of *biscuit* to that in Arabic بسكويٲ baskawi:t. Here, participants tried to reproduce a written form that matches the source loan spelling though it is actually pronounced as *biskit* in English. As to *sandwich*, the influencing factor can be seen in choosing loan-form 1 (ساندوتش sa:ndwitʃ) with the English sound tʃ which does not have an equivalent in Arabic but can be approximated in pause form pronunciations a consonant cluster tʃ (broken by a vowel in full pronunciation form). Participants attempted to preserve English tʃ in Arabic spelling to be close to the English spelling form. Loans under loan-form 2 (*diploma*, *dolphin*, *catalog*) demonstrate violation of native analogical patterns informed partly by the English orthography. In the native loan-forms of دبليوم diblu:m 'diploma' and دولفين du:lfi:n 'dolphin', o > u:, i > i: in *dolphin* and a > a: in كتالوج kata:ludʒ 'catalog' and, accordingly, these loans developed non-analogical patterns: fiʔl:ul, fuʔli:l, faʔa:lu:l whereas the analogical patterns in Arabic are faʔlu:l, fiʔli:l, and faʔa:li:l respectively. This lengthening of English vowels when Arabicized may be referred to orthography; in that Arabic short sounds (a, i, u) do not have orthographic correspondents like long ones but rather are represented by diacritics.

¹ Both loan-forms are Arabic equivalents, however, loan-form 1 is closer to the English spelling than loan-form 2.

Thus, non-long vowels in loans are mapped to their long counterparts (a:, i:, u:) when nativizing them and are reflected in the written forms. What happens here is that by maintaining English spelling in Arabic, the loan-forms become close to English spelling but native analogical patterns are compromised. The influence of English spelling on Arabic is also present in loan compounds, although such loans are generally limited in Arabic. The loan entries in table 4 are characterized by analyzability where loans under loan-form 1 were selected better be written in two parts whereas the other ones under loan-form 2 were preferred be written as single part words. ?ais.kri:m ايس كريم 'ice cream' and tɪfɪ:z.kayk تشيز كيك 'cheesecake' demonstrate the effect of orthography in which bilinguals perceived these loan compounds as two-unit words. They are more aware to the analyzability of these words in English reflecting them in Arabic spelling. In hamburger همبرجر 'hamburger' and mi:krufu:n ميكرفون 'microphone', bilinguals selected a single word written form following English spelling. Table 4 below illustrates loan compound entries (most selected form in bold) along with the number of responses of the most selected form:

Table 4: loan compound with multiple written forms

Loanword	Loan-form 1	Loan-form 2	# of responses
ice cream	?ais.kri:m ايس كريم	?iskir:m ايسكريم	10
hamburger	ham.burgar هام برجر	hamburger همبرجر	12
microphone	mækru.fu:n مكرفون	mi:krufu:n ميكرفون	11
cheesecake	tɪfɪ:z.kayk تشيز كيك	ʃi:zkayk شيز كيك	12

The frequency of occurrence affects the choice of a particular spelling. The majority of loans forms (except كيبنة kabi:na 'cabin') selected by participants strongly correlate with the most frequent forms in arTenTen12as illustrated in table 4. The frequency effect develops during early encounter with loans and reflects speakers' behavior in taking on a common form. Institutional efforts such as language academies, educational policies, and media contribute to the promotion of particular loan-forms over others which, in return, affect the frequency of spelling as well as pronunciation. It is generally accepted that media most often uses standard Arabic and, thus, the choice of written forms are institutionally informed; in accordance with language policy. Table 4 below is an example of frequency effect in media which agree with the most preferred loan-forms selected by participants.

Table 4: loan forms frequency in Arabic Web Corpus

Loanword	Loan-form 1	Freq	Loan-form 2	Freq
diploma	diblu:ma دبلوما	921	diblu:m دبلوم	42,807
cabin	kabi:na كيبنة	1,309	ka:bi:na كابينة	3,328
biscuit	baskawi:t بسكويت	27	basku:t بسكوت	5
dolphin	du:lfi:n دولفين	874	du:lfi:n دولفين	3,066
catalog	kataludʒ كتالوج	184	kata:ludʒ كتالوج	2,261
sandwich	sa:ndwiʃ ساندوتش	1,133	sandawif سندوش	1
ice cream	?ais.kri:m ايس كريم	2,120	?iskir:m ايسكريم	520
hamburger	ham.burgar هام برجر	126	hamburger همبرجر	239
microphone	mækru.fu:n مكرفون	285	mi:krufu:n ميكرفون	1,498
cheesecake	tɪfɪ:z.kayk تشيز كيك	32	ʃi:zkayk شيز كيك	5

It can be argued that variation in loan-form reflects the speaker's unsuccessful attempt to be faithful to both languages during adaptation where orthography emerges as a challenge that disrupts perception resulting in competing loan-forms. The orthography of the source language seems to play an important role in adaptation and this accounts for bilinguals attempt to replicate the spelling of English loans in Arabic forms even though they are pronounced differently. This role is motivated by a close conformity between phonemes and alphabets in standard Arabic (see Gordon, 1970, 193–97). It is further enhanced by literacy where reading and writing are developed in standard Arabic and, therefore, orthography is important in Arabic for religious and cultural purposes unlike indigenous American communities for instance. The debate on the educational reform excluded orthography which becomes no longer a concern (Haeri, 2000). In other words, the system of orthography in standard Arabic is difficult to counter and change. The choice of orthographic forms in accordance with standard Arabic justifies speakers' reliance on standard spelling, since spoken varieties are unwritten and will be ridiculed if used. In fact, variation in written forms of loans within standard Arabic is just one aspect of a larger scope of variation between standard Arabic and other spoken varieties.

Al-Wer (1997) pointed out that the use of particular form from CA over others from Arabic spoken varieties undergoes some sort of ideology. She referred this ideology to the religious and cultural attachment of Arabs to CA which failed to provide a standard form of communication across Arab world. She argued that this bias against modern spoken varieties, which are restricted in education, calls for recodifying standard Arabic since "... linguistic variation and change in Arabic involves interplay between local varieties and emerging regional standards." (p: 262). This may explain the conformity of the majority of loan-forms chosen by participants to those in standard Arabic as seen in table 4. It can be claimed, then, that participants' choice of loan-forms is affected by their education that enforces standard Arabic patterns especially in writing.

The findings of this study contribute to the existing literature on loanword adaptation by introducing the role of orthography at the phonological and morphological levels with different degrees of influence beside the sociolinguistic implication of such role. The results agree with Vendelin and Peperkamp (2006), Daland, M., and Kim (2015) in that transformation of foreign sounds seems to predicate on a phonetic nature and influenced by orthographic representations of source language. The effect of orthography is more salient with vowels due to the limited number of vowels in Arabic along with short vowels being invisible in written form. Although the findings did not clearly answer the questions of this study, orthography stems as an influencing factor in adaptation in both single and compound loans. It can be claimed, then, that the adaptation process is more phonetically oriented following (Silverman, 1992; Peperkamp & Dupoux, 2003; Peperkamp, 2004; Kabak and Idsardi, 2007) where bilinguals map foreign sounds to their native correspondents based on similarity and partly influenced by orthography. This can be seen in vowels of source loans (a, i, u) that are lengthened as well as in consonants where $tʃ > tʃ$ and $p > b$ in *cheese cake* and *cabin*.

Orthography of the source loans affects the morphology of loan-forms by initiating illegal analogical patterns in L1. The violations to analogical patterns in $دبلوم$ *diblu:m* 'diploma' and $دولفين$ *du:lfi:n* 'dolphin' are partly triggered by orthography. If the spelling effect is to be set aside, these loans can be analogical as $infaʔlu:l > dablu:m$ and $fiʔli:l > dilfi:n$. As to the sociolinguistic factors of loans incorporation, bilinguals' choice of spelling is biased conforming to the forms of standard Arabic even though they are aware that some loans are pronounced differently in English, i.e. *baskawi:t* for *biskit* 'biscuit'. In the absence of written forms for spoken varieties, bilinguals seem to adhere to the standard forms regardless of whether they are adequate. Obviously, these primary findings are promising and suggest further empirical work and statistical examination over a larger sample of population and number of loanwords to understand more about the effect of orthography on loanword adaptation.

6. Conclusions

Loanwords are a multifaceted phenomenon that can be studied from different perspectives. Although phonological and morphological theories of loanword adaptation are substantial, there are other extra linguistic factors that intervene in their integration and use. The findings from this study suggest a significant role of loans orthography when adapting them to Arabic. The approximation of the spelling of loans in Arabic to English is unjustifiable affected native phonological and morphological patterns. The choice of orthographic form correlates with standard Arabic even if it is non-analogical. Variation in loan-forms should not be seen as free but rather as a consequence of bilinguals' unsuccessful attempts to transfer foreign sounds based on similarity of their L1 in a phonetically oriented process. This paper calls for experimental research in search of orthography and psycholinguistic roles in loanword adaptation in Arabic to enrich existing literature with more insights.

References

- Al-Qinai, J. (2000) Morphophonemic's of Loanwords in Arabic. *Studies in the Linguistics Sciences*, 30, 2.
- Al-Saidat, E. (2011). English Loanwords in Jordanian Arabic: Gender and Number Assignment. *Language Forum*, 37(1), 59- 72
- Al-Wer, E. (1997). Arabic between reality and ideology. *International Journal of Applied Linguistics*, 7(2), 251-265.
- Bahumaid, S. (2015). Lexical Borrowing: The Case of English Loanwords in Hadhrami Arabic. *International Journal of Language and Linguistics* 2, 6, 13-24.
- Belinkov, Y., Habash, N., Kilgarriff, A., Ordan, N., Roth, R., & Suchomel, V. (2013). arTenTen12: Anew, vast corpus for Arabic. In *Second workshop on Arabic Corpus Linguistics (WACL-2)*, July 22.
- Broselow, Ellen. 2004. Language contact phonology: Richness of the stimulus, poverty of the base. In Keir Moulton and Matthew Wolf, eds., *Proceedings of NELS 34*. Amherst, MA: GLSA, 1-21
- Campbell, Lyle (2004). *Historical Linguistics: an Introduction* (2nd edition). Edinburgh: Edinburgh University Press, and Cambridge, MA: MIT Press.
- Campbell, Lyle and Verónica Grondona. 2010. Who speaks what to whom? Multilingualism and language choice in Misi3n La Paz. *Language in Society* 39:617
- Chang, C. B. (2008). Phonetics vs. phonology in loanword adaptation: Revisiting the role of the bilingual.
- Comrie, Bernard. 2000. Language contact, lexical borrowing, and semantic fields. In Gilbers, Dicky & Nerbonne, John & Schaeken, Jos (eds.), *Languages in Contact (Studies in Slavic and General Linguistics 28)*, 73–86. Amsterdam: Rodopi.
- Daland, R., Oh, M., & Kim, S. (2015). When in doubt, read the instructions: Orthographic effects in loanword adaptation. *Lingua*, 159, 70–92. doi:10.1016/j.lingua.2015.03.002
- Detey, S., Nespoulous, J.-L., 2008. Can orthography influence second language syllabic segmentation? Japanese epenthetic vowels and French consonantal clusters. *Lingua* 118, 66–81
- Dohlus, K. (2005) Phonetics or phonology: Asymmetries in loanword adaptations—French and German mid front rounded vowels in Japanese. *ZAS Papers in Linguistics* 42: 117—135
- Epps, Patience and Kristine Stenzel. 2013. Upper Rio Negro: cultural and linguistic interaction in Northwestern Amazonia. Rio de Janeiro: Funai.
- Gordon, Cyrus H. 1970. The accidental invention of the phonemic alphabet. *Journal of Near Eastern Studies* 29(3):193–97.
- Haeri, N. (2000). Form and ideology: Arabic sociolinguistics and beyond. *Annual Review of Anthropology*, 29(1), 61-87.
- Hafez, O, M. (1996) Phonological and Morphological Integration of Loanwords into Egyptian Arabic, *Égypte/Monde arabe*, Première série, Les langues en Égypte
- Hamdi, S. (2017). Number and Gender Assignment to Loanwords in Arabic: Implications from Varieties. *International Journal of Linguistics*.
- Haspelmath, M. (2009). II. Lexical borrowing: Concepts and issues. *Loanwords in the World's Languages*. doi:10.1515/9783110218442.35
- Haugen, E. (1950). The Analysis of Linguistic Borrowing. *Language*, 26(2), 210. doi:10.2307/410058
- Hussein, R. F. (1999), Code-Alteration among Arab College Students. *World Englishes*, 18: 281–289. doi:10.1111/1467-971X.00141
- It3, J., Mester, A., 1995. The core-periphery structure of the lexicon and constraints on reranking. In: Beckman, J., et al. (Eds.), *Papers in Optimality Theory*. Amherst, University of Massachusetts, Amherst, pp. 181–210.
- Jacobs, H., Gussenhoven, C., 2000. Loan phonology: perception, salience, the lexicon, and OT. In: Dekkers, J., et al.
- Kabak, B., Idsardi, W.J., 2007. Speech perception is not isomorphic to phonology: the case of perceptual epenthesis. *Lang. Speech* 50, 23-52

- Kang, Y. (2003). Perceptual similarity in loanword adaptation: English postvocalic word-final stops in Korean. *Phonology*, 20(02), 219–273. doi:10.1017/s0952675703004524
- Kenstowicz, M., & Suchato, A. (2006). Issues in loanword adaptation: A case study from Thai. *Lingua*, 116(7), 921–949. doi:10.1016/j.lingua.2005.05.006
- Kenstowicz, M., 2001/2004. The role of perception in loanword phonology. *Linguistique Africaine* 20; *Studies in African Linguistics* 32, 95–112.
- LaCharité, D., & Paradis, C. (2005). Category Preservation and Proximity versus Phonetic Approximation in Loanword Adaptation. *Linguistic Inquiry*, 36(2), 223–258. doi:10.1162/0024389053710666
- Paradis, C., & LaCharité, D. (1997). Preservation and minimality in loanword adaptation. *Journal of Linguistics*, 33(2), 379–430. doi:10.1017/s0022226797006786
- Peperkamp, S. (2004). A Psycholinguistic Theory of Loanword Adaptations. Annual Meeting of the Berkeley Linguistics Society, 30(1), 341. doi:10.3765/bls.v30i1.919
- Peperkamp, S., & Dupoux, E. (2003, August). Reinterpreting loanword adaptations: the role of perception. In *Proceedings of the 15th international congress of phonetic sciences* (Vol. 367, p. 370).
- Poplack, S., & Sankoff, D. (1984). Borrowing: the synchrony of integration. *Linguistics*, 22(1), 99-136.
- Poplack, S., Pousada, A., & Sankoff, D. (1982). Competing influences on gender assignment: Variable process, stable outcome. *Lingua*, 57(1), 1–28. doi:10.1016/0024-3841(82)90068-7
- Poplack, S., Sankoff, D., & Miller, C. (1988). The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics*, 26(1). doi:10.1515/ling.1988.26.1.47
- Poplack, S., Sayahi, L., Mourad, N., & Dion, N. (2015). An Exception to the Rule? Lone French Nouns in unisian Arabic. *University of Pennsylvania Working Papers in Linguistics*, 21(2), 178-186.
- Sankoff, David, Poplack, Shana and Vanniarajan, Swathi. 1990. The case of the nonce loanin Tamil. *Language variation and change* 2:71–101.
- Shinohara, S., 1997. *Analyse phonologique de l'adaptation japonaise de mots e'trangers*. The`se de doctorat, Universite Paris III [ROA-243].
- Silverman, D. (1992). Multiple scansion in loanword phonology: evidence from Cantonese. *Phonology*, 9(02), 289. doi:10.1017/s0952675700001627
- Steriade, D. (2001). The phonology of perceptibility effects: the P-map and its consequences for constraint organization. Ms., UCLA.
- Thomason, Sarah Grey & Kaufman, Terrence. 1988. *Language contact, realization, and genetic linguistics*. Berkeley: University of California Press
- Vendelin, I., & Peperkamp, S. (2006). The influence of orthography on loanword adaptations. *Lingua*, 116(7), 996–1007. doi:10.1016/j.lingua.2005.07.005
- Yip, M. 1993. Cantonese loanword phonology and Optimality Theory. *Journal of East Asian Linguistics* 2: 261-291