Number and Gender Assignment to loanwords in Arabic: Implications from Varieties

Sami Hamdi (Corresponding author)

Dept. of Language, literacy, and Sociocultural Studies, University of New Mexico, Hokona, 140, MSC 05 3040

University of New Mexico, Albuquerque, New Mexico 87131-1231, English Language Center, Jazan University, P.O. Box 114 Jazan, KSA, Postal Code 45142, Saudi Arabia

Tel: 1-505-610-8760   E-mail: sami_max2002@hotmail.com

Abstract

Number and gender are two of the core grammatical categories in Arabic. The assignment of number and gender to foreign words is an area of conflict between MSA and other Arabic varieties. This paper investigates the factors that stand behind the seemingly irregularity of number and gender assignment in Arabic. It appears that speakers follow a form standardized by MSA or enforce another form following their dialects and community conventions. This variation in number and gender assignment to loans gives rise to multiple competing forms that may not be recognized by MSA or some other varieties of Arabic. Yet, the findings demonstrate consistency in assigning number and gender to loans by applying native patterns motivated by frequency of use and the semantics of the referents.

Keywords: Loanwords, Arabic, Number assignment, Gender assignment

1. Introduction

Loanwords are the words that are transferred from one language and integrated into another language. Loanwords have been one of the interesting linguistic phenomena that challenge the grammatical system of the recipient language during adaptation. The degree of loanwords integration in a language is characterized by different levels of linguistic changes to meet the grammatical system. The phonological and morphological levels are the most relevant ones during adaptation.

Arabic, like many other languages, receives a lot of loanwords and adapts them into its grammatical system. Typologically speaking, structural constraints are thought to account for the higher borrowability of nouns comparing to verbs. Tadmor (2009,p. 63) associates structural constraints with whether the recipient language is isolating or synthetic. He puts it that morphosyntactic adaptation of borrowed verbs tends to be lower when the recipient language is more isolating, whereas higher level of morphosyntactic adaptation is needed when the language is synthetic. This generalization applies to Semitic languages such as Arabic. Since Arabic is a highly inflectional language, loanwords are marked for a set of grammatical categories such as person, mode, gender, and number. For nominal loans, four categories are involved: number, gender, definiteness, and case.

During morphological adaptation of foreign words, number and gender assignment emerge as a conflict area between Standard Arabic and other Arabic varieties. However, this conflict shouldn't be seen completely as ad hoc especially when taking into consideration that Arabic doesn't have a neutral gender such as English.

In this paper, our focus is on number and gender assignment of nominal loans in Arabic. Although there are consistent grammatical patterns in assigning number and gender in Arabic, differences emerge when it comes to Arabic varieties. For instance, *cabins* can have kabina:t and kaba:ʔin as well as *workshop* that can have weraʃa:t and weraʃ as plural forms.

2. Number and Gender in MSA

In this section, we brief inflectional process concerning number and gender in MSA using loanwords in the examples. For certain cases in MSA nouns may have special forms to mark number and gender, yet here we show the standard patterns which are relevant to nominal loans. Number assignment in MSA inflects for singular, dual, and plural. The singular represents the base form which is unmarked as in /bank/, whereas the dual and plural forms are marked by suffixes. The dual form has the suffix /-a:ni/ in the nominative case and /-ayni/ in the accusative and genitive cases as shown below:

(nominative)	(accusative)	(genitive)
bank-**a:ni**	bank-**ayni**	bank-**ayni**
bank- two	bank- two	bank- two
'two banks'	'two banks'	'two banks'

The plural is formed based on whether the referent is animate or inanimate. There are three pluralization methods: *sound masculine plural* to code human male referents (or a mixed

group of males and females) using suffixes /-uːn/ for nominative and /-iːn/ in the accusative and genitive case. The *broken plural* is unpredictable applying irregular vowel change within the word itself. Finally, the suffix /-aːt/, the *sound feminine* plural marker, is added to the word final position. The sound plural method indicates a base form that remains intact when pluralized. Table 1 below demonstrates how loanwords are inflected for number:

Table 1. Nominal adaptation of loanwords

Loanword	Broken pl.	Sound Msc.pl.	Sound Fem.pl
cabin	kabaːʔin		kabinaːt
workshop	weraʃ		weraʃaːt
radar		#	radaːrt
telephone		#	telifuːnaːt
musician		musiːqiy-uːn/iːn	musiːqiy-aːt (f)
comedian		kumidiy- uːn/iːn	kumidiy-aːt (f)

From the above table, it can be observed that the example loanwords have inanimate referents, except *musician* and *comedian*. They are inflected for plurality through the sound feminine plural form. This tendency is very common in Arabic (Ryding, 2005) and it applies to a wide range of nouns whose referents are human or nonhuman. However, it should be emphasized here that the use of the feminine sound plural does not necessarily corresponds to the original gender of the loanword. A loanword can be masculine in the single form; i.e. *radaːr* and feminine in the plural form *radaːrat* for 'radars'. In other words, the sound feminine plural seems to be highly productive strategy and is applied for generating plural forms. The broken plural is unpredictable applying vowels insertion within the word; what can be called ablaut. However, some loans can have sound plural forms along with the broken plural forms. Overall, the broken plural is the last resort when sound plural doesn't apply.

There are only two genders in Arabic: masculine and feminine which are usually overtly marked in the adjectives, nouns, pronouns, and verbs. The assignment of gender is regular when the referent is animate, however, gender assignment is semantically arbitrary when the referent is inanimate. The masculine gender represents the basic category whereas the suffix /-taː/ and its variants /-a/ and /-ha/, in pause form, code the feminine form as in /daktuːr/ and /daktuːr**a**/ for male and female doctor respectively.

3. Previous Studies

Loanword adaptation in Arabic has been researched from different perspectives focusing on the varieties of Arabic more than Modern Standard Arabic (MSA). Smeaton (1973) studied the morphological integration of loanwords in Hasawi Arabic and describes three stages: Except for phonemic adaption, loans preserve their original forms and if pluralized only /-aːt/ (a sound feminine plural marker) is added word finally. Then, they can be shortened or expanded to match possible native patterns. Finally, loans are nativized morphologically if

they are nouns having internal pluralization and, if verbs, they should be able to inflect for all verbs forms. These modifications are made for harmony with native root system morphology.

Al-Qinai (2002) described the common word derivation processes in MSA to incorporate foreign words in Arabic such as clipping, compounding, remodeling, derivation and inflection. However, most of the loanwords in Al-Qinai's are outdated and adapted by early philologists. Hafez (1996) examined synchronically the different degrees of integration of loans in Egyptian English comparing them to their original source language. She identifies three levels of integration: derivation of a fully-fledged paradigm form, addition of a feminine suffix, and plural inflection of nouns.

In Hadrami Arabic, Bahumaid (2015) points out that number assignment agrees with native rules where masculine loans ending in consonants are pluralized by the sound feminine marker /-aːt/ whereas for loans ending in vowels a double of semi vowels –yy are inserted before the feminine marker as in loːri – loːriyyaːt for lorry – lorries. Gender assignment follows native patterns where loanwords with final consonants in the singular forms are often masculine (e.g., iryal 'aerial'; breːk 'brake'). The feminine form is coded by the feminine markers –ah or –ih in   shoːt**ih** 'shot'; garm**ah** 'garment'.

Poplack, Sayahi, Mourad, and Dion (2015) studied French lone nouns in Tunisian Arabic (TA) and pointed out that about 91% of lone French nouns in TA were used following the native grammar patterns. They reported that there was a limited number of lone French nouns that are often pluralized using the sound feminine marker /-**ɛ**ːt/. TA tends to employ the broken plural marker more than the regular sound markers. Although this zero-coding of plural seems inconsistent with the native pattern on the surface level, it conforms to the other native morphological rules such as when the referents are more than ten in quantity or when their equivalents are singulars as in (e.g., dollar[ø] and lunettes 'glasses').

Alsaidat (2011) examined gender and number assignment of English loanwords in Jordanian Arabic (JA). He points out that gender assignment in Jordanian Arabic is based on the phonetic ending and on the noun signification; that is whether the form has a masculine/ feminine marker or it denotes a particular gender in spite of the form. Number assignment in JA follows MSA in the singular form while the dual is marked using a neutral marker /-ein/ and its variant /-hein/. For pluralization, the marker /-aːt/ and its variant /-haːt/ are employed.

4. The Current Study

It can be noticed from the previous works that researchers paid greater attention to loanword adaptation in several varieties of Arabic with limited reference to MSA. In this study, a set of borrowed words in Arabic varieties are examined in comparison to MSA. The governing factors behind number and gender assignment are investigated. Loanwords are used to investigate number and gender assignment as areas of conflict between MSA and other Arabic varieties. The terms foreign words, loans, and loanwords are used interchangeably. The loanwords entries were provided by the researcher and validated for their number and gender patterns through Almaany Online Dictionary.

5. Arabic and Variation

Arabic is a Semitic language that exhibits a rich content of loanwords; mainly nominal. 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 (the language of religious teaching and early literary works). It is used in writing, news, and formal communication across Arabs world. Other varieties of Arabic spread as regional dialects over the Arab world with different degrees of intelligibility cross dialectally. These varieties are spoken and used mainly in every day communication and informal settings. There are grammatical and stylistic aspects that distinguish MSA from other varieties such as word order, lexical choice, and morphological derivations. Actually, the characterization of variation in Arabic is a matter of ongoing debate that is out of the scope of the present paper. However, our approach is variational and limited to the comparison of number and gender assignment of nominal loans between MSA and other varieties.

5.1 Number Assignment in Varieties

This section demonstrates number and gender assignment in a set of Arabic varieties in comparison to MSA. The dual form seems to be less considerably observed as a variation among the different varieties of Arabic. Yet, it is worth mentioning that the suffix /-ayn/ and its variants cross-dialectal are the most frequent regardless of the case ending.

The plural formation is an area of conflict in Arabic showing variation, thus it is emphasized when discussing number assignment. Table 2 below lists a set of loanwords and their plural forms. Only loans that exist in MSA and other varieties were cited[1]. The variety that uses a different plural form is marked next to the loan-form as an example, but the form is not necessarily used only by that variety.

Table 2. Plural forms of loans cross dialectally

Loanwords	MSA pl.	Broken pl. in varieties
lorry	luːriyyaːt	lawaːri (EA)
radio	raːdjuhaːt	radaːwi (EA)
gallon	dʒaluːnaːt	galaneIn (JA) or dʒawaliːn in other varieties
villa	fillaːt	v/fillal (EA)
blouse	bluːsaːt	balaːjIz (JA)
cigarette	sidʒar /-aːt/-ʔir/-jIr/	sagaːjIr (JA)
machine	maːkinaːt	maːkinaːt/ maːkaːɖin

Loanwords in Table 2 reflect a competing inflectional strategy of the plural form. MSA applies the sound plural form as the norm whereas other varieties enforce the broken plural form even if there is a standardized available form. This conflict area of plural formation can be referred to: frequency of usage; in that speakers use the most common form in their variety

---

[1] Some loanwords are used in a specific variety but not present in MSA

whether they are aware of the standard form or not. When speakers use MSA, they would most likely use the form following MSA pattern. The case of 'cigarette' in table 2 further supports the frequency factor where MSA has three possible plural forms. Here, the speakers of varieties are believed to apply the form that agrees with their native pattern when using their variety and MSA contexts. Furthermore, the use of the broken plural in varieties is characterized by phonotactics where it is convenient for speakers to apply their native patterns.

The broken plural is unpredictable and, thus, avoidable whenever possible in MSA while for the varieties it is flexible and common. The tendency to use the broken plural form in varieties can't be seen only as preference or free variation. It is a matter of enforcing native patterns in line with the conventional norms of the speech community.

5.2 Gender Assignment in Varieties

The varieties of Arabic show different gender assignment of certain loans from MSA. The long held belief that gender is arbitrarily assigned might need be revised at least with loanwords case. Assigning gender to loanwords seems to be motivated by applying the nearest equivalent form based on semantic analogy (in MSA/variety) to the loanword as in table 3.

Table 3. Gender variation based on equivalent

Loanwords	G- MSA	G- varieties	Motivation
balloon	M	F	
internet	F	M	Semantic
centimeter	M	F	analogy

Another factor behind differences in gender assignment is motivated by the sense of the referents. In table 4, a set of loanwords are presented having different or multiple genders in MSA and other varieties:

Table 4. Gender variation based on sense

Loanwords	G- MSA	G- varieties	Sense 1- MSC	Sense 2- Fem	Motivation
cream	M	F	ointment	whipped cream	
saloon	M/F	M/F	business for haircuts or beauty/home for regular meetings of writers or artists	living room/a hall for exhibition of art	Different sense
flash	M	F	a sudden burst of light	flash memory	

In table 3 and 4, the difference in gender assignment between MSA and other varieties is motivated by two criteria: semantic analogy based on the nearest equivalent form or the particular sense of the referents. A loanword is assigned a gender based on its nearest equivalent in MSA as in *balloon*; matching the nearest native/translated form /mintˤaːd/ which is masculine. When a loanword is assigned both the masculine and feminine genders,

the difference is based on the gender of the referents in a specific sense. For instance, *cream* is masculine when used in the sense of *ointment* while feminine when referring to *whipped cream*. This characterization of the distribution of gender seems consistent in MSA and this might be due to institutional efforts to regularize foreign words in MSA. Gender assignment in varieties, that don't follow MSA, can be attributed to speakers applying the nearest native equivalents in their dialects regardless of whether it conforms to MSA.

The notion in Arabic literature that gender assignment to inanimate referents is semantically arbitrary is not strongly justified. It is interesting to note that some loanwords denote a profession as a shared semantic feature and are associated with a particular gender. For example, loans like *cashier*, *captain*, *chauffeur*, and *supervisor* are used to refer to professions dominated by males more than females. Likewise, other loans such as nurse /na:rsa:/ are used in JA Alsaidat (2011) to refer to female nurse since nursing as a profession is associated with females. Thus, there should be a semantic level/frame involved in assigning gender to foreign words urging speakers to treat them as masculine or feminine. Obviously, such a generalization, to be more valid, has to be empirically supported over a large number of loanwords.

6. Discussion

The areas of conflict between MSA and other varieties concerning foreign words are the pluralization method during number assignment and with the particular sense of referents during gender assignment. MSA prioritizes regularizing foreign words applying a sound plural form and avoiding, whenever possible, broken plural form since it is irregular. Gender is assigned to loans with respect to their nearest equivalents or Arabized form in MSA. On the other hand, broken plural in Arabic varieties is highly employed where multiple forms can be used with internal changes; inserting vowels within words. When assigning gender, varieties apply the local native form regardless of whether it agrees with MSA. Furthermore, there is no neutral gender in both MSA and varieties which may allow variation to take place especially when foreign words are originally assigned neutral gender in the source language.

There are different factors that account for variation such as the degree of bilingualism and education as well as the socio economic status of speakers. Ibrahim (1973) describes some factors that operate during gender assignment of loans such as semantic content, homophones between the languages in question, and if the language has grammatical gender. However, the frequency of use seems the most important factor when it comes to variation. Speakers use the most frequent forms in their speech community. When a native form is competing with another form (from MSA/other dialects), speakers are likely to enforce their native dialectal forms following community conventions. Poplack and Sankoff (1984) explained that "conflict in gender assignment is a transitory stage on the route to assimilation of certain loanwords, and tends to disappear as frequency of use and phonological integration increase" and this reflects a period of instability of gender assignment to loans before they take on a specific gender.

Another factor is triggered by context where MSA forms or the most frequent forms cross dialectally are chosen over a native form. For instance, *machine* can have three plural forms:

ma:kina:t, ma:ka:ʔin, and makan. The first form is the standard MSA sound plural form whereas the other two are broken plural forms. The majority of speakers would use either the first form (consistent with MSA) or the second form, which is common, cross dialectally but less likely to use the third form. This situation can be seen when speakers from different dialects are to use words that have multiple inflectional methods and feel forced to use a common form (other than their native) to communicate with speakers from other dialects.

However, bias is also possible where speakers prefer their native form and resist using another possible form even if it is more frequent. Lastly, analyzability is also an essential grammatical factor in Arabic. Foreign words in Arabic are treated as solid stems that can't be broken down into units. For MSA, this treatment may preserve the aspect of foreignness of loans in comparison to native words or equivalents. In varieties, some regular patterns are avoided for simplicity and this is apparent in using multiple forms of broken plural which are zero-suffix.

Poplack, Pousada, and Sankoff (1982) pointed out that the variationist approach on gender assignment is language specific rather than universal. They emphasized the factors behind the initial assignment of gender. This explanation applies to Arabic case in both gender and number assignment. The perceived ad hoc assignment of the grammatical number and gender of loans with inanimate referents can be understood and resolved when looking at speakers' behavior during first encounter or use of foreign words. The various factors behind number and gender assignment are to be investigated in varieties more than the standard dialect.

The classification or identification of the nature of variation between MSA and other Arabic varieties is a complicated task. Yet, grammatical categories such as case and number can be viewed as sources of variation. Studying variation of different linguistic aspects in Arabic should not be confused with the regularity of patterns in comparison to the standard forms, since variants are consistent with the native grammar of the dialects. Owens (2001, p. 453) pointed out two dimensions that describe the incorporation of items from standard Arabic to other native varieties: borrowing lexemes that introduce new ideas and the use of multiple forms as controlled by stylistic and contextual reasons. Foreign words, especially cultural terms, have fairly stable lexical stems and when integrated from MSA into dialects, the variation concerns inflectional patterns such as number and gender assignment. Most of varieties speakers are aware of the standard forms but switch to less standard forms for particular factors mentioned above.

7. Conclusion

Assigning number and grammatical gender to foreign words is a conflict area between MSA and other Arabic dialects. During lexical borrowing, number and gender assignment are necessary grammatical categories for agreement. For number assignment, the unmarked form is masculine in the singular form which is the basic. Loan nouns are overtly marked to code gender based on the semantic aspect of animacy of referents as well as the equivalent and the particular sense of the referents in use. The sound feminine plural form is a common mechanism to generate plural despite the referent's gender. MSA assigns number and gender to foreign words following a standard pattern. Although other varieties recognize multiple

forms, they prefer to use their own native local pattern of their speech community to other linguistic patterns from MSA. Resolving variation in number and gender assignment can be achieved by finding the nearest native equivalent and the sense of referents in use. A set of factors can explain the different distribution of number and gender, yet some factors are speakers' or dialect specific.

References

Almaany. (n.d). [Online] Available: http://www.almaany.com/

Al-Qinai, J. (2000) Morphophonemics 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

Bahumaid, S. (2015). Lexical Borrowing: The Case of English Loanwords in Hadhrami Arabic. *International Journal of Language and Linguistics, 2*(6), 13-24.

Cadora, F. J., & Smeation, B. H. (1977). Lexical Expansion due to Technical Changeas Illustrated by the Arabic of al-Hasa, Saudi Arabia. *Journal of the American Oriental Society, 97*(2), 210. http://dx.doi.org/10.2307/599022

Hafez, O, M. (2008) Phonological and Morphological Integration of Loanwords into Egyptian Arabic, Égypte/Monde arabe, Première série, Les langues en Égypte

Ibrahim, M. H. (1973). Grammatical Gender. doi:10.1515/9783110905397

Owens, J. (2001). Arabic Sociolinguistics. *Arabica, 48*(4), 419–469. http://dx.doi.org/10.1163/157005801323163816

POPLACK, S., & SANKOFF, D. (1984). Borrowing: the synchrony of integration. Linguistics, 22(1). doi:10.1515/ling.1984.22.1.99

Poplack, S., Pousada, A., & Sankoff, D. (1982). Competing influences on gender assignment: Variable process, stable outcome. *Lingua, 57*(1), 1–28. http://dx.doi.org/10.1016/0024-3841(82)90068-7

Poplack, S., Sayahi, L., Mourad, N., & Dion, N. (2015). An Exception to the Rule? Lone French Nouns in Tunisian Arabic. *University of Pennsylvania Working Papers in Linguistics, 21*(2), 178-186.

Ryding, Karin C. (2005). A Reference Grammar of Modern Standard Arabic. Campridge. Campridge University Press. http://dx.doi.org/10.1017/cbo9780511486975

Tadmor, U. (2009). III. Loanwords in the world's languages: Findings and results. Loanwords in the World's Languages. http://dx.doi.org/10.1515/9783110218442.55

Copyright Disclaimer