

University at Albany, State University of New York

## Scholars Archive

---

Languages, Literatures and Cultures Faculty  
Scholarship

Languages, Literatures & Cultures

---

2012

# A Quantitative Analysis of Code-switching in the Arabic-Romance Kharjas

Lofti Sayahi

*University at Albany, State University of New York, sayahi@albany.edu*

Juan Antonio Thomas

*Utica College, juantomas329@yahoo.com*

The University at Albany community has made this article openly available.

**Please share** how this access benefits you.

Follow this and additional works at: [https://scholarsarchive.library.albany.edu/cas\\_llc\\_scholar](https://scholarsarchive.library.albany.edu/cas_llc_scholar)

---

### Recommended Citation

Sayahi, Lofti and Thomas, Juan Antonio, "A Quantitative Analysis of Code-switching in the Arabic-Romance Kharjas" (2012). *Languages, Literatures and Cultures Faculty Scholarship*. 41.

[https://scholarsarchive.library.albany.edu/cas\\_llc\\_scholar/41](https://scholarsarchive.library.albany.edu/cas_llc_scholar/41)



This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

This Article is brought to you for free and open access by the Languages, Literatures & Cultures at Scholars Archive. It has been accepted for inclusion in Languages, Literatures and Cultures Faculty Scholarship by an authorized administrator of Scholars Archive.

Please see [Terms of Use](#). For more information, please contact [scholarsarchive@albany.edu](mailto:scholarsarchive@albany.edu).



## A Quantitative Analysis of Code-switching in the Arabic-Romance *Kharjas*

Juan Antonio Thomas<sup>a</sup> and Lotfi Sayahi<sup>b,\*</sup>

a) Utica College

juantomas329@yahoo.com

b) State University of New York at Albany

sayahi@albany.edu

---

### Abstract

This paper studies code-switching between Andalusian Arabic and Romance in the *kharjas*, the closing verses of the *muwashshahaat* poems. These poems, dating from the 11<sup>th</sup> to the 14<sup>th</sup> centuries, were composed in Classical Arabic, while the *kharjas* were written in two languages of Al-Andalus: Andalusian Arabic and Romance. The purpose is to investigate to what degree the structural aspect of code-switching in the *kharjas* conforms to descriptions in the current literature on code-switching in bilingual communities and what that tells us about the degree of bilingualism in Al-Andalus. The 43 *kharjas* (Corriente, 2008) present a total of 104 code-switches: 82 intra-sentential, 13 word-internal and 9 inter-sentential. The base language in the majority of cases is Romance: 73 % of the switches occurred from Romance to Arabic. Cross-tabulations of the direction of the switch, lexical category of the switched parts and what immediately precedes and follows them show statistically significant relationships, indicating that the code-switches found in this corpus are not the result of a random process of language mixing resulting in “*an outlandish and deliberately unsophisticated patois*” (Monroe, 1974:31). A study of the intra-sentential code-switches also contributes to an explanation of the behavior of the Arabic definite article, *al-* and its allomorphs, in Arabic loanwords.

### Keywords

code-switching; languages in contact; language change; bilingualism; Andalusian Arabic; Romance

## 1. Introduction

The language contact situation in Al-Andalus was particularly rich. The Muslim armies did not bring one sole variety of Arabic with them but rather Classical Arabic along with vernacular Arabic, dependent upon the origin of

---

\* We are very grateful to two anonymous reviewers for their suggestions and corrections. All errors remain our own.

the speaker, all in a diglossic relationship. The Arabic spoken in Al-Andalus developed into what is called Andalusian Arabic, studied extensively by Corriente (1977, 1992, 1999). Andalusian Arabic is considered a member of the Western Arabic dialect group that includes North African Arabic vernaculars and Maltese. This group presents substratal influence from Berber, a language that, given the composition of Muslim armies and the different North African dynasties that governed Al-Andalus, was in contact with Andalusian Arabic. Mozarabs, who were Christians living in Al-Andalus, spoke a variety of Romance in the early Middle Ages that showed characteristics quite different from the variety of vulgar Latin which would develop into modern Castilian. In the present work, we will use the term Romance to refer to the variety of Romance spoken in Al-Andalus (Galmés de Fuentes 1983, 1984), which was in contact with Andalusian Arabic and is the Romance found in the *kharjas*.

Castilian Spanish, even though it developed outside of Al-Andalus, was in contact with Mozarab Romance, as well as with vernacular varieties of Arabic. The fruit of such prolonged contact led to the incorporation of approximately 4000 words, including toponyms,<sup>1</sup> based on Arabic etyma (Lapesa 1981: 133). Indeed, Arabic is a considerable lexifying source for Spanish. On the other hand, while the influence of Spanish on Classical Arabic is nonexistent, the North African dialects have some traces from this language as a result of the introduction of Spanish in different periods since the 15<sup>th</sup> century (Sayahi, 2011, 2012). An example is the case of Northern Moroccan Arabic which has hundreds of Spanish loanwords as a result of its contact with Spanish before and during the Spanish Protectorate in the 20<sup>th</sup> century (Heath 1989; Ghailani 1997; Sayahi 2005).

After so many years, it may seem impossible to analyze the nature of the complex contact situation in Al-Andalus. However, poets from that period have left us a genre of Hispano-Arabic poetry known as *muwashshah*, composed between the 11<sup>th</sup> to the 14<sup>th</sup> centuries. The *muwashshahaat* (plural of *muwashshah*) were written entirely in Arabic script and composed in Classical Arabic, but the ending, known as the *kharja* /*xar.ʒal* 'exit; closure', was composed in Romance and Andalusian Arabic, with code-switching between the two languages. Monroe (1974: 29) puts it this way: "*The most important part of the poem was the kharja or "exit" which was frequently in Romance or in*

---

<sup>1</sup> Lapesa (1981: 133) explains that 850 Spanish words have clearly identified Arabic etyma and that another 780 are derivatives. There are approximately one thousand toponyms of certain Arabic origin and about 500 of probable Arabic origin. In the present work, we have isolated 1290 loanwords of Arabic origin (*vide infra*).

vernacular Arabic [...] rang[ing] from ironic refusal to crude obscenity, all in an outlandish and deliberately unsophisticated patois”.

*Kharja* A22 (Corriente 2008: 250) is reproduced in (1). Uppercase and lowercase letters represent Romance and Arabic elements, respectively. The translations and glosses into English are by the authors. News of the *kharjas*

- (1) *ALBA DÍYA ÉSTA DÍYA, DÍYA D + alšāṣara ḥāqqa,*  
‘A HAPPY DAY IS TODAY, DAY OF the most righteous one,’<sup>2</sup>

*BEŠTIRÉY MEW almudabbáj wanišúq arrúmḥa šáqqa*  
I WILL WEAR MY BROCADE SUIT and I will break my lance.’

first appeared in Stern’s pioneering articles in 1948 and 1949, followed by that of García Gómez (1952), which contained many of the Romance *kharjas* found in the *Uddat al-Jalīs*, one of the anthologies of *muwashshahaat*. García Gómez further spread his views in his 1965 volume, including the belief that the *kharjas* form part of a primitive Romance lyric tradition in the Iberian Peninsula, cultivated by the Romance speakers in Al-Andalus who maintained that oral legacy even during the Muslim domination. Solà-Solé’s anthology of *kharjas* (1973) offers alternative interpretations to those of García Gómez although relies on many of the latter’s texts. Alan Jones published the first detailed paleographic and linguistic analysis of the *kharjas* in 1988. More recently, researchers, such as Zwartjes (1995), have argued that while the *kharjas* do have some extra Arabic realities, they were composed “completely according to Arabic tradition: historically, prosodically, thematically and stylistically” (Zwartjes, 1995: 248). Corriente’s contributions (1997, 2004, 2008, 2009), based on Jones (1988) and his own work on the Andalusian Arabic dialect (1977, 1992, 1999), offer more refined readings of the *kharjas*, and we base this investigation on the forty-three *kharjas* in Corriente (2008).

Although many authors have commented on the code-switching in the *kharjas*, there has been no systematic study of the phenomenon to support their judgments. Jones (1988: 11) believes that “in many cases the switches from one language to the other and back again are not normal or convincing if judged by what has been learned in recent years about language switching”. Similarly, Monroe (1974: 31) classifies the *kharjas* as “an outlandish and deliberately unsophisticated patois”. Whinnom (1982: 15) even denies that code-switching exists in the *kharjas* and suggests that what we consider to be Romance might be loanwords in Andalusian Arabic. Sempere Martínez, on the other hand,

<sup>2</sup> Corriente (2009) translates *DÍYA D + al šāṣara ḥāqqa* as ‘summer solstice’.

takes exception to those judgments: “*Es una incongruencia proponer una deficiencia gramatical en un texto que, por otra parte, constituye una elaborada composición lingüística. El lenguaje de las jarchas bilingües- y el de las monolingües también- no puede ser fortuito.*” (Sempere Martínez 1998: 646). [It is an incongruence to propose a grammatical deficiency in a text that, on the other hand, constitutes an elaborate linguistic composition. The language of the bilingual *kharjas*- and also that of the monolingual ones- cannot be fortuitous. (translated by authors)]

Sempere Martínez shows how recent scholarship on code-switching can provide more convincing interpretations of some earlier readings by Solà-Solà and García Gómez.

The objective of the present study is to determine the frequency, types, directionality of the switch, lexical categories of the switched parts and structure of the examples of code-switches (CS) found in the *kharjas*. This information will help determine whether or not the code-switching observed in the *kharjas* is structurally different from what is found in the literature about code-switching. Such data should also provide more information about the type of bilingualism in Al-Andalus.

We will also see that the code-switching data sheds some light onto how Romance may have incorporated the Arabic definite article (*al*- and its allomorphs) in loanwords. We will analyze the behavior of the Arabic definite article in the *kharjas* data as well as in a corpus of Arabic loanwords extracted from the *Diccionario de la Lengua Española* (1995). We consider the study of the Arabic article to be especially important as the *kharjas* offer a way to examine directly the process of the integration (or not) of the Arabic article into loanwords. A long-standing misconception regarding Arabic loan nouns in Romance is that they systematically conserve *al*- in the first syllable. As Don Quixote explains to Sancho how to identify Arabic loanwords, Cervantes echoes this commonly held belief: “... *tanto por el al primero como por el i en que acaban, son conocidos por arábigos* [.. as much as for the initial *al* as for the *i* in which they end, they are recognized as Arabic] (Cervantes, 731)”.

The *kharjas* are literary creations. We are not dealing with spontaneously produced, natural speech samples. However, owing to the high frequency of vocatives, short quotes and the themes of the *kharjas*, which are basically street talk in a much more informal register than the *muwashshahaat*, the language use in the *kharjas* most probably approximates that of the lower class of Al-Andalus. Most of the *kharjas* are anonymous, although we know the names of many of the poets who composed the *muwashshahaat*.

Several studies have argued for the role of code-switching as a mechanism for language contact and change (Thomason 2001; Myers-Scotton 2002; Backus 2005). In general, this investigation will show that code-switching in the *kharjas* bears witness to a high degree of bilingualism in Al-Andalus that led to intense language contact and, by consequence, lexical borrowing.

## 2. Methodology

Even though investigators such as Whinnom (1982) and Jones (1988) have commented on the arbitrary nature of the code-switches found in the *kharjas*, we propose a method here to determine whether or not those claims are founded. We characterized the directionality of the languages switched after considering the totality of the *kharja*. All *kharjas* cited herein correspond to Corriente's (2008) designations for the Arabic series (A). In *kharja* A21 (see 2), the noun phrase *alḥabībE* is an example of Romance to Arabic intra-sentential code-switching. The phrase *ya māmma*, present throughout the corpus, contains a Romance to Arabic intra-sentential code-switch.

- (2) *ya MĀMMA, MEW alḥabībE FÓY ŠE EN + E(D) NON TORNÁDE,*

“Oh, Mother, my lover went away and he does not return,

*GĀR KĒ FAREYO, ya MĀMMA? E(D) NIN BEYJĒLLO LEŠÁDE (A21)*

Say, what should I do, Mother? And he doesn't leave me even a little kiss.”

The sample consists of data obtained from the forty-three *kharjas* interpreted by Corriente (2008). Earlier readings are found in Corriente (1997, 2004). Code-switches were counted and were classified according to the direction of the switch (Arabic to Romance, or Romance to Arabic), the type of switch (inter-sentential, intra-sentential and word-internal) and the lexical category of the switched item, as well as the lexical categories of the elements both preceding and following the switched item.

Given that the determination of the parameters conditioning the code-switches forms the heart of this study, an additional illustration of the codification is presented in (3, 4, 5). This is *kharja* A3.

- (3) *VĒT EN! wúč YA TENRAD*

“go away! What a face (nerve) he must have then” (A3)

Example 3 is a case of intra-sentential CS from Romance to Arabic. The switched item is the Arabic *wúč*, a noun meaning ‘face’ used here to indicate ‘daring’ or ‘nerve’. The preceding element is the Romance locative pronoun

EN, but since the switched item, *wúć*, is the direct object of the Romance verb *TENRAD*, the switch is considered to be intra-sentential.

- (4) *κ indár xáLES KÉRED*  
“he wants to alert the uncles” (A3)

Example 4 is another case of a Romance to Arabic intra-sentential CS. The switched item is the Arabic verb *indár*, followed by the Arabic *xáLES*. Preceding the switch is the Romance conjunction *κ*. This is considered Romance to Arabic since the phrase is embedded between the Romance elements, *κ* and *KÉRED*.

- (5) *xáLES KÉRED*  
uncles he wants (A3)

Example 5 is a word-internal Arabic to Romance CS. Preceding the CS, is the Arabic noun *xál* “uncle”. The CS element is the Romance plural morpheme *-ES*, followed by the Romance verb *KÉRED*.

The switched element, as well as the elements before and after it, were assigned to fourteen lexical categories: noun, pronoun, nominal morpheme, verb, verbal morpheme, adjective, adverb, determinative-article, determinative-demonstrative, determinative-possessive, preposition, conjunction, interjection, or no element. Data summaries and statistical analyses were performed with the Statistical Package for Social Science (S.P.S.S.). Pearson chi-square probabilities less than .050 are considered significant. In a separate analysis, 1290 loanwords from Arabic were isolated from the *Diccionario de la lengua española* (1995) and were classified according to the presence of the definite article *al-*, its assimilation, or its absence.

### 3. Results and Discussion

#### 3.1. The types of code-switches in the corpus

The total number of switches in the *kharjas* is 104, an average of 2.4 switches per *kharja*. Table 1 summarizes the frequencies and percentages of the two types of language shift possible, Romance to Arabic and vice versa. Table 1 also gives the frequencies of the three types of code-switching found in the *kharjas*: intra-sentential, inter-sentential and word-internal.

The number of code-switches found in this corpus of the *kharjas* is abundant: 104 cases present in 43 *kharjas*, which is high given that a typical *kharja* consists of two to four verses. Since Romance to Arabic shifts dominate

**Table 1** Frequencies and types of language shift in the *kharjas*

	intra-sentential	word-internal	inter-sentential	total
Romance > Arabic	61	9	6	76 (73 %)
Arabic > Romance	21	4	3	28 (27 %)
total	82 (79 %)	13 (12 %)	9 (9 %)	104

(73 %) with respect to Arabic to Romance (27 %) and *kharjas* A16 and A18 were entirely composed in Romance, the latter tends to be the base language in the Arabic-Romance *kharjas*.

Intra-sentential switches were the most frequent type, observed at a frequency of 79 %. Inter-sentential switches were found at a rate of 9 %. The fact that the *kharjas* are not extended conversations explains the scarcity of this type. The intra and inter-sentential code-switches observed here are not categorically structurally different from what is found in the literature on code-switching.

Interestingly, a type of switch, which we considered word-internal, was observed in 12 % of the total switches. An example was given in (5) and additional cases are shown in (6), where the code-switched element is underlined. These are word-internal code-switches from Romance to Arabic, where inflection/ derivation was performed on Arabic bases; for example, the verbal root *liqá* with the Romance morpheme *-R*, and the noun and adjective bases *ljumm* and *šaqr* with Romance diminutive suffixes *-ELLO/A*.

(6) GARRÉ(D)ME ÓB liqáRTE  
“TELL ME WHERE TO find YOU” (A1)

SO ljummÉLLA šaqrÉLLA  
“UNDER locks LITTLE blond LITTLE “ (A14)

Word-internal code-switching has been flagged as ungrammatical by some researchers (cf Poplack's (1980) free morpheme constraint), who most likely would consider these to be nonce or spontaneous borrowings. Regardless of their names, it has been shown that in bilingual communities lexical insertions in the form of spontaneous borrowings are fairly common (Heath 1989; Owens 2002). These cases imply a high level of bilingualism as opposed to established loanwords that can be used by monolingual speakers (Poplack, Sankoff and Miller 1988: 52). Given that some of these examples consist of the same borrowed morpheme, such as the Romance diminutive *-ELLO/A*,



these switches indicate the level of access the authors and audience must have had to both languages in order to fully understand them.

This analysis provides evidence that the poets had competence in both languages. The high degree of code-switching suggests the existence of extended bilingualism since the poets would not compose or incorporate bilingual *kharjas* if their audience could not understand them. Given that the main part of the *muwashshah* was written in Classical Arabic, it is most likely that they were Arabic-dominant bilingual speakers, but the use of Romance, presumably their weaker language, suggests that the poets were advanced bilinguals. Code-switching at the end of a poem written in Classical Arabic between two vernacular systems, Andalusian Arabic and Romance, indicates that the contact was between these two and not between Classical Arabic and Romance, consistent with Mufwene's contention that: "Overall, we learn that vernaculars compete with vernaculars and *lingua francas* with *lingua francas*" (Mufwene 2007: 84). Even if the author of a *muwashshah* used a *kharja* originally composed by a different poet, still, the high informality of the themes in the *kharjas* in a way reflects the linguistic behavior that must have been in use among the authors and audience of these poems. These are speakers who use Classical Arabic, a language that is only acquired through formal instruction, but who also use the two vernaculars that are available to them in their daily life and even code-switch between the two.

### 3.2. *Lexical category and type of code-switching*

Table 2 summarizes the cross-tabulation of the lexical category by the type of the CS, as well as by the direction of the switch, that is, whether the code-switch proceeds from Romance to Arabic (R>A) or from Arabic to Romance (A> R). The cross-tabulations of Romance to Arabic switches, which constitute the majority of code-switches (76 cases), by the three types of CS yield a Pearson chi-square probability ( $p$ ) = .002. The most frequently involved lexical categories were: 17 nouns; 11 cases of *ya*, the Arabic vocative particle; and 24 articles, indicating that most such CS occurred within a noun phrase. The cross tabulation of the code-switched lexical category and type of code-switch shows that these two variables have a non-random relationship: 83 % ( $p$  = .002) of the Romance to Arabic inter-sentential code-switches were initiated by the verb. Nouns and Arabic determinants, specifically definite articles, dominate the intra-sentential switches; 37 out of 59 Romance to Arabic intra-sentential CS (61 %) and 4 out of 9

**Table 2** Cross tabulation of the lexical category of the code-switched part

direction		type of code-switch			total
		intra-sentential	word-internal	inter-sentential	
R > A	noun	15	2	0	17
	conj	1	0	1	2
	det-art	22	2	0	24
	inter	2	0	0	2
	verb	4	3	5	13
	prep	2	0	0	2
	adv	3	0	0	3
	adj	1	2	0	3
	voc	11	0	0	11
A > R	total	61	9	6	76
	noun	8	0	1	9
	conj	1	0	0	1
	pronoun	1	0	0	1
	det-poss	1	0	0	1
	det-dem	1	0	0	1
	nominal morpheme	0	4	0	4
	verb	7	0	1	8
	adv	2	0	1	3
total	21	4	3	28	

word-internal CS (44 %) involve nouns. Because of the high frequency of nouns and definite articles involved in the code-switches analyzed here, we are able to observe how the Arabic article behaved in Arabic insertions into Romance discourse.

Table 3 summarizes the lexical categories of the words before, after and at each case of intra-sentential code-switching, although the latter is included in complete detail in table 2. For the Romance to Arabic code-switches, the cross-tabulations of both the lexical category of the preceding element with the type of code-switch as well as the part of speech of the following element with the type of code-switch showed a non-random relationship, as attested by the Pearson chi-square probabilities.

Therefore, this analysis shows that the CS syntax has a significant relationship with the lexical category of the switched parts, demonstrating an underlying pattern to the syntax, rather than a haphazard mixing of Romance and Arabic elements.

**Table 3** Occurrences of intra-sentential CS

	preceding element	code-switched element	following element
Romance > Arabic (61 cases)	noun: 8 verb: 12 det-poss: 5 prep: 13 pron: 9 none: 5 the rest < 5 p=.017	noun: 15 verb: 4 art: 22 voc: 11  the rest < 4 p=.002	noun: 40 verb: 4 art: 4 adv: 5  the rest ≤ 4 p = .002
Arabic > Romance (21 cases)	noun: 8 adv: 3 none: 3 the rest ≤ 2	noun: 8 verb: 7  the rest ≤ 2 p=.035	noun: 5 verb: 4 adv: 4 the rest ≤ 2

### 3.3. The Arabic definite article and double determinants

In order to relate the above findings about code-switching in the *kharjas* to the process of the introduction of the Arabic article into Romance loanwords, the structure of the noun phrases in the code-switching junctures are relevant, including allomorphs of the Arabic definite article. The main allomorph of the article is *al-*. However, a total reverse coronal assimilation oftentimes modifies the consonant to that of the initial consonant of the noun, for example, *arraqib* ‘the watcher/guardian.’ In this case, the /l/ is assimilated to the following /t/ as it does with all coronal consonants in this context.

Before accounting for the behavior of the article in the majority of the noun phrases, we would like to report a less frequently observed usage in the *kharjas*, but one where Arabic and Romance grammar rules mix to allow two syntactical contexts where a double determinant is expressed, the definite article and either a demonstrative or possessive. Since this poetry was written by and for bilingual Arabic-Romance speakers, it is unlikely that the Arabic definite article would have been expressed for no syntactic motivation; bilinguals cannot ignore the article. Three Arabic nouns, prefaced by the Arabic article, were also preceded by a Romance possessive: *MEW alhabib*, in A8 and A21 and *SEW arraqib* in A28. While medieval Castilian syntax allowed definite articles

to precede possessive adjectives, such as “*el mio fiel vassallo, la mi muger*” (Penny 1991:126), just as some modern Romances still do, the possessive in the *kharjas* precedes the definite article. Nevertheless, because the Arabic definite article is agglutinated to the nouns, this possessive construction observed in the *kharjas* may be considered a calqued metathesis of Romance syntax. Likewise, the two Arabic nouns, again with the attached definite article and prefaced by a demonstrative: *EST arraqīf* “this rogue” and *EST alḥaraki* “this troublemaker”, both in A10, could be calques of the Arabic syntax. This construction requires the demonstrative adjectives to precede definite nouns in order to create phrases, rather than sentences such as *EST raqīf* “this is a rogue” and *EST ḥaraki* “this is a troublemaker”.

#### 3.4. The code-switching junctures in the noun phrases

In the Romance to Arabic intra-sentential CS, of the 15 nouns that initiated the code-switch, none was preceded by a definite article, and since 13 of them were not preceded by any determinant, they are considered bare nouns. Of the 22 articles that initiated a code-switch, all 22 were followed by a noun and 10 of the articles were preceded by a preposition. This indicates that the Arabic nouns prefixed by the Arabic article were introduced into Romance discourse via a prepositional phrase at a higher frequency (10/ 22 or 45 %) than were nouns with no article (3/15 or 20 %). After all of the 11 intra-sentential code-switches initiated by the Arabic vocative *ya*, a noun followed, sometimes a Romance noun, such as *QORATSONE*, *MATRE*, *MAMMA*, other times an Arabic noun, such as *sidi*, *ḥabibi*, *saḥḥāra*. No noun introduced after a code-switched *ya* was accompanied by the definite article; hence, they were bare. In all cases, the vocative phrase can be considered a parenthetical element to the discourse, but it was not counted as an inter-sentential CS.

Summing up the 3 types of Romance to Arabic code-switches involving nouns, of the 45 Arabic nouns (see table 4), 23 are bare nouns (51 %) and 22 nouns (49 %) have the Arabic definite article. Thirteen have the non-assimilated allomorph *al-* and 9, the assimilated one. Ten of the 22 nouns with definite articles were preceded by prepositions (45 %). Prepositional phrases are one of the main structures that introduce Arabic nouns with the articles into the *kharjas*. The bare nouns included 5 examples of the Arabic vocative *ya* followed by an Arabic noun and another 6 cases of a vocative without *ya*, for 11 vocatives (46 %). Vocative phrases tend to introduce bare nouns into the *kharjas*.

**Table 4** Arabic single noun code-switches in the *kharjas*

nouns that contain <i>al-</i>	nouns that contain the assimilated form	bare nouns
<i>alšāšiqā</i> “the beloved”	<i>aḏḏāmmaš</i> “the embraces”	<i>sīdī</i> (2) “my lord”
<i>alḥābib</i> (2) “the lover”	<i>aššāmaš</i> “the beauty marks”	<i>abrāhim</i> “Abraham”
<i>alḥarakī</i> “the troublemaker”	<i>arraḡīb</i> (2) “the watcher”	<i>abulqāsīm</i> “Abdulqasim”
<i>alfalakÉ</i> “the stars”	<i>aššūhdÉ</i> “the honey”	<i>abulḥajjāj</i> “Abulhajjaj”
<i>alšānšara</i> “the most righteous one”	<i>aššārṭi</i> “the condition”	<i>amīri</i> “my prince”
<i>alġilāla</i> “the shirt”	<i>arraḡīf</i> “the rogue”	<i>allāhi</i> “My God”
<i>allāzmaš</i> “marks from biting”	<i>annādmÉ</i> “the necklace”	<i>ammi</i> “my mother”
<i>alwāʿid</i> “the promises”	<i>assamrÉLLO</i> “the little dark one”	<i>saḥḥāra</i> “sorceress”
<i>alḥujāj</i> “the excuses”		<i>ḥābb almulūk</i> “cherry”
<i>alšiqdÉ</i> “the pearl necklace”		<i>xalÉLLO</i> (2) “little thief”
<i>alġulām</i> “the boy”		<i>ḥābib</i> “lover”
<i>almudabbāj</i> “the brocade jacket”		<i>ḥābibi</i> (2) “my lover”
		<i>lġummÉLLA</i> (2) “little locks”
		<i>bāssÉ</i> “kisses”
		<i>xāmri</i> “syrup”
		<i>bilāsi</i> “nothing”
		<i>ṭurrahÁYRA</i> “trifle”
		<i>wāyAš</i> “sighs”
		<i>wūč</i> “face”

### 3.5. The Arabic article in Spanish

We also measured the number of Spanish nouns of Arabic origin that were adapted with and without the definite article. Using the electronic version of the *Diccionario de la lengua española* (1995) that allows searches by origin, all entries that were tagged as being of Arabic origin were classified according to the behavior of the Arabic article. The interesting result is that, using this method, half the entries that the Spanish Academy considers to be of Arabic origin do not in fact conserve the Arabic article. The other half is divided between words that keep the article intact (31 %) and those that present its assimilated form (19 %). These results bear a striking similarity to the behavior of the article in code-switched Arabic nouns in the *kharjas*. Table 5 summarizes the data gathered from the *Diccionario de la lengua española*.

**Table 5** Arabic loanwords recorded in the *Diccionario de la lengua española*

number of loanwords	loanwords with <i>al-</i>	loanwords with the assimilated form	loanwords without the article
1290	400 (31 %)	242 (19 %)	648 (50 %)

### 3.6. Why the Arabic article is or is not incorporated in loanwords

Solà-Solé (1968) estimates the number of Arabic loanwords in Spanish to be between 850-1000; 60 % contained a form of the article. Bramon (1987) analyzed 229 cases of Arabic loanwords present in both Spanish and Catalan; 70 % contained a form of the article. In our study of the *Diccionario de la lengua española*, we have counted 1290 loanwords where 50 % are expressed with the article and 50 % without the article. Of the Arabic nouns involved in the Andalusian Romance to Arabic code-switches in the *kharjas*, 51 % were bare nouns and 49 % had the Arabic article.

In Arabic the definite article is agglutinated and is required for each element of the noun phrase. It is invariable across number and gender and undergoes a reverse assimilation before coronal consonants, which comprise 14 out of 28 sounds in Arabic. For example, the adapted loanwords in Spanish, *almohada* ‘pillow’ and *aceite* ‘oil’, contain the article in the non-assimilated and assimilated allomorphs, respectively.

There are several theories that attempt to explain why the definite article is retained in the Arabic loanwords. Steiger as cited by Solà-Solé (1968: 280) invokes communicative reasons for the presence of the article: “[...] *se debe su presencia o su ausencia a las exigencias del pensamiento o de la comunicación*. [...its presence or absence is due to the demands of thought or of communication...]” Solà-Solé (1968: 280) further expounds on the presentative function of the article: “*El artículo árabe al- desempeña no solo una función determinativa y, a veces de resultas de su propio origen, demostrativa, sino además, y por encima de todo, una función que podríamos llamar PRESENTATIVA o ENFÁTICA* [The Arabic article *al-* carries out not only a determinative function, and, at times as a result of its very origin, demonstrative, but also, and above all, a function that we could call PRESENTATIVE or EMPHATIC]”. Steiger (1932) and Elcock (1960) believe that a Berber substrate influence is responsible for the predominance of the article in Andalusian Arabic. Corriente (1999: 61), referring to the Arabic spoken by speakers of Berber, says: “...*cuando*

*pasaron a Alandalús, difundirían entre los nativos un tipo de árabe criollizado o pidginizado, con un uso abusivo del artículo...[...when they passed to Al-Andalus, a type of creolized or pidgeonized Arabic with an abusive use of the article would diffuse among the natives...]*". Yet, Noll (1996) cites bilingualism and prestige as the main factors for the maintenance of the article.

Based on the code-switching data found in the *kharjas*, we would like to suggest a new mechanism that explains both why and why not the Arabic definite article is incorporated in the loanwords. Many Arabic loanwords were introduced into Romance by code-switching. If the syntactic structure allows for the frequent expression of the article, it stands more chances to be incorporated in the loanword. While we do not insist that prepositional phrases are the only structures capable of carrying along an article, in this corpus, they were frequent vehicles for delivering nouns with articles. Bare nouns tended to be introduced in vocative phrases. Hence, Arabic loanwords are accompanied by the definite article in those syntactic contexts where it is appropriate to express the definite article. The nature of the article did not lead to its borrowing directly. The syntactic context in which it was more frequently expressed may, hence, contribute to explain its retention.

This code-switching mechanism has the ability to explain the existence of doublets, that is, two words that come from the same etymon, one with the article incorporated and the other without. As we see in table 4, a noun like *ḥabīb/alḥabīb* appears both with and without the article, as a function of the syntactic context in which it is found.

### 3.7 *The kharjas: literary creations or language of the street?*

A final word is warranted regarding the fact that the *kharjas* are literary creations and not spontaneous speech samples. The *muwashshahaat* could only be composed after the poet had spent years learning Classical Arabic. Although the *kharja* is also a literary creation, the fact that it has Romance elements breaks with the strict codes of an Arabic formal register. The high frequency of vocative phrases that we have noted is yet more evidence that the poets were trying to imitate natural speech, as are the topics of the 'conversations' in the *kharjas*, such as the curses of the jilted young women pining for their absent lovers or calling upon their mothers for advice. It seems that the *kharjas* were intended to reflect at least the practices of the street, if not the very language.

#### 4 Conclusion

The analysis of code-switching junctures in the *kharjas* shows that these are related to the lexical category of the switched part, demonstrating that their syntactic structure is not randomly constructed. The dominance of intra-sentential code-switches, the existence of word-internal code-switches and the use of Romance, in an otherwise highly formal poem composed in Classical Arabic, show that the poets were Arabic-dominant advanced bilinguals. In order to understand these works, extended bilingualism must have characterized at least the audience of the *kharjas*, if not a good part of the society of Al-Andalus.

Since many of the code-switches occur in noun phrases, the *kharjas* give us a context to observe how the article is incorporated (or not) in Arabic loanwords. Indeed, code-switching is most likely to have played a role in bringing Arabic loanwords into Romance and the possibility of retaining the article where allowed by the syntactic structure. As we have shown, code-switching from Romance to Arabic shows usage of the article according to Arabic rules. We can see from this study that Arabic-dominant speakers, such as the poets who composed the *muwashshahaat* and *kharjas* and their audiences, played a role in introducing Arabic items into Romance that could be subsequently picked up by Romance-dominant speakers. Borrowing into Romance was not a process that concerned Romance-dominant Mozarabs only but it also must have involved bilingual Arabic speakers who were agents of linguistic influence in both directions.

#### References

- Backus, Ad. 2005. Codeswitching and language change: One thing leads to another? *International Journal of Bilingualism* 9: 307-340.
- Bramon, Dolors. 1987. Aglutinación y deglutinación del artículo en los arabismos del castellano y del catalán. *Vox Romanica* 46: 138-180.
- Cervantes Saavedra, Miguel de. 2000. *El Ingenioso Hidalgo Don Quijote de la Mancha*. Barcelona: Editorial Optima.
- Corriente, Federico. 1977. *A grammatical sketch of the Spanish Arabic dialect bundle*. Madrid: Instituto Hispano-Árabe de Cultura.
- Corriente, Federico. 1992. *Árabe andalusí y lenguas romances*. Madrid: Editorial MAPFRE.
- Corriente, Federico. 1997. *Poesía dialectal árabe y romance en Alandalús: cejeles y xarajāt de muwaššahāt*. Madrid: Gredos.
- Corriente, Federico. 1999. *Diccionario de arabismos y voces afines en iberorromance*. Madrid: Gredos.



- Corriente, Federico. 2004. *Twenty-Seven Muwashshahaat and One Zajal by Ibn al-Arabi of Murcia (1165-1240)*, (en colaboración con E. Emery). London: School of Oriental and African Studies.
- Corriente, Federico. 2008. *Romania árabe: tres cuestiones básicas: arabismos, "mozárabes" y "jarchas"*. Madrid: Editorial Trotta.
- Corriente, Federico. 2009. The kharjas: an updated survey of theories, texts, and their interpretation. *Romance Philology*, 63 109-129.
- Elcock, William Denis. 1960. *The Romance languages*. London: Faber & Faber.
- Galmés de Fuentes, Álvaro. 1983. *Dialectología mozárabe*. Madrid : Gredos.
- Galmés de Fuentes, Álvaro. 1994. *Las jarchas mozárabes: forma y significado*. Barcelona: Crítica.
- García Gómez, Emilio. 1952. Veinticuatro kharjas romances en muwajjahas árabes (Ms. G.S. Colin). *Al-Andalus* 17: 57-127.
- García Gómez, Emilio. 1965. *Las jarchas romances de la serie árabe en su marco*. Madrid: Sociedad de Estudios y Publicaciones.
- Ghailani, Abdellatif. 1997. *El Español hablado en el Norte de Marruecos*. Doctoral dissertation, Universidad de Granada.
- Heath, Jeffrey. 1989. *From code-switching to borrowing: Foreign and diglossic mixing in Moroccan Arabic*. London: Kegan Paul International.
- Jones, Alan. 1988. *Romance Kharjas in Andalusian Arabic Muwashshah Poetry*. London: Ithaca Press.
- Lapesa, Rafael. 1981. *Historia de la lengua española*. Madrid: Gredos.
- Monroe, James T. 1974. *Hispano-Arabic poetry: a student anthology*. Berkeley: University of California Press.
- Mufwene, Salikoko. 2007. Population movements and contacts: Competition, selection, and language evolution. *Journal of Language Contact* 1: 63-91.
- Myers-Scotton, Carol. 2002. *Contact linguistics: bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- Noll, Volker. 1996. Der arabische Artikel al und das Iberoromanische. In Jens Lüdtke (ed.), *Romania arabica. Festschrift für R. Kontzi zum 70 Geburtstag*. 299-313. Tübingen: Narr.
- Noll, Volker. 2006. La aglutinación del artículo árabe al en el léxico español. In Rafael Arnold and Jutta Langenbacher-Liebott (eds.), *Cosmos léxico. Contribuciones a la lexicología y la lexicografía hispánicas XIV* 35-49. Peter Lang Verlagsgruppe: Frankfurt/M.
- Owens, Jonathan. 2002. Processing the world piece by piece: iconicity, lexical insertion and possessives in Nigerian Arabic codeswitching. *Language Variation and Change* 14: 173-209.
- Penny, Ralph. 1991. *A History of the Spanish Language*. Cambridge: Cambridge University Press.
- Poplack, Shana. 1980. Sometimes I'll start a sentence in Spanish Y TERMINO EN ESPAÑOL: Toward a typology of code-switching. *Linguistics*, 18: 581-618.
- Poplack, Shana, Sankoff, David, and Chris Miller. 1988. The social correlates and linguistic processes of lexical borrowing and assimilation. *Linguistics* 26:47-104.
- Real Academia Española. 1995. *Diccionario de la Lengua Española* (21 ed.). Madrid: Espasa Calpe.
- Solà-Solé, Josep Maria. 1968. El artículo al- en los arabismos del iberorrománico. *Romance Philology* 21: 275-285.
- Solà-Solé, Josep Maria. 1973. *Corpus de poesia mozárabe(las kharjas andalusies)*. Barcelona: Hispam.
- Sayahi, Lotfi. 2005. Phonological adaptation of Spanish loanwords in Northern Moroccan Arabic. *University of Pennsylvania Working Papers in Linguistics* 11 (1): 253-263.

- Sayahi, Lotfi. 2011. Spanish in Contact with Arabic. In *Handbook of Spanish Sociolinguistics* 473-490. Malden, MA: Wiley-Blackwell.
- Sayahi, Lotfi. 2012. Contacto y préstamo léxico: el elemento español en el árabe. *Revista Internacional de Lingüística Iberoamericana* 18: 85-100.
- Sempere Martínez, Juan A. 1998. Agramaticalidad e improbabilidad de alternancia de códigos en las jarchas. In Claudio García Turza, Fabián González Bachiller, and Javier Mangado Martínez (eds.), *Actas del IV congreso internacional de historia de la lengua española* vol II, 643-651. Logroño: Universidad de La Rioja Servicio de Publicaciones.
- Steiger, Arnald. 1932. Contribución a la fonética del hispano-árabe y de los arabismos en el ibero-románico y el siciliano. *Revista de Filología Española*, 17. Madrid: Casa Editorial Hernando.
- Stern, Samuel. 1948. Les vers finaux en espagnol dans les muwaffaḥs hispano-hébraïques: une contribution à l'histoire du muwaffaḥ et à l'étude du vieux dialect espagnol 'mozárabe'. *Al-Andalus* 13: 299-348.
- Stern, Samuel. 1949. Un muwaffaḥ arabe avec terminaison espagnole. *Al-Andalus* 14: 214-228.
- Thomason, Sarah Grey. 2001. *Language Contact*. Washington D.C.: Georgetown University Press.
- Whinnom, Keith. 1982. The MAMMA of the Kharjas or some Doubts Concerning Arabists and Romanists. *La Corónica* 11: 11-17.
- Zwartjes, Otto Jeroen. 1995. *The Andalusian Xarja-s: Poetry at the Crossroads of Two Systems?* Nijmegen: University of Nijmegen.