LONE ENGLISH-ORIGIN NOUNS IN ARABIC: CODESWITCHES OR BORROWINGS?*

Eiman Mustafawi
University of Ottawa

1. Introduction

The categorization of lone words from one language in the discourse of another as *igroub* in (1) has a bearing on different topics in bilingualism, such as, the characterization of borrowing, the phenomenon of codeswitching, and most importantly the constraints on codeswitching.

(1) fi Canada day δε:le:n l-igroub illi ywagf-u:n w iyγann-u:n
in these the-group that stand-they and sing-they
w iy-rigS-u:n.
and dance-they
‘On Canada day, this is the group that stands, sings and dances’

Imperical research has shown that such single items constitute the majority of the other language material in most bilingual discourse, so grouping them with the wrong category may obscure the patterns of behavior of the true members of that category. For example, if lone words are categorized with codeswitches, their patterns of behavior may skew the patterns of behavior of the true codeswitches, which gives rise to theories of codeswitching which account poorly for the data (Ghafar-Samar & Meechan 1998, p. 206).

Therefore, it is important to keep the status of lone words ambiguous until their patterns of behavior show similarity, to either established loanwords (borrowings) or unambiguous codeswitches (CSs).

There are three different views in the field with respect to lone words. The first as reflected in work of Mahoutian (1993), Eliasson (1990), and Myers-Scotton (1992; 1993), does not distinguish between borrowing and codeswitching and attribute them to the same mechanism.

The second view considers any single word from a donor language that is not an established loanword in the recipient language to be a codeswitch. This

*I would like to thank Shana Poplack and Marjory Meechan for their insightful comments on an earlier version of this paper. I would also like to thank the members of the sociolinguistics meeting group for their helpful comments.

Actes de l’ACL 2002/ 2002 CLA Proceedings
view is reflected in (e.g. Bentahila and Davies 1991; Bokamba 1988; Eliasson 1994, among others), and in most of the bilingual literature involving Arabic (c.f. Al-Mansour 1998; Nortier 1990; Myers-Scotton 1995; Mohamed 1989; Sallo 1994; Hussein & Shorrab 1993; Atawneh 1992; Bentahila and Davies 1983, 1991). Some of these authors claim that for a lexical item to be considered a borrowing it must fill a lexical gap in the recipient language (Atawneh 1992, p. 238).

The third view with respect to lone words holds that borrowing and codeswitching are different mechanisms. That is, in codeswitching the integrity of the grammar of both the donor and the recipient languages is respected, while in borrowing only the integrity of the grammar of the recipient language needs to be respected (Poplack 1993, Poplack and Meechan 1995, 1998, Poplack, Wheeler and Westwood 1989, Budzhak-Jones 1998, Samar and Meechan 1998, Eze 1998, Turpin 1998). Also, borrowings do no necessarily “fulfill lexical needs” in the recipient language (Poplack, Sankoff, and Miller 1988). Poplack (1993, p. 256) proposes the nonce-borrowing hypothesis, according to which, single words and compounds can be borrowed “momentarily”. Accordingly, lone words are best characterized as borrowings, even though their distribution across the community is currently limited.

**Nonce Borrowing Hypothesis:** “Nonce borrowing differs from codeswitching, and resembles established borrowing in all but its extralinguistic characteristics of recurrence and diffusion” (Poplack and Meechan 1998, p. 137).

Applying a quantitative analysis and the comparative variationist methodology instanciated in Poplack and Meechan (1998), I test the validity of the nonce borrowing hypothesis. It was found that the category of nouns provides an ideal context for such a study, partly because nouns out-number any other category in bilingual data and because they provide a considerable number of conflict sites between the two languages.

By using the appropriate tools for this kind of study, I will show that the lone English-origin nouns in Arabic/English bilingual discourse were adopted into the Arabic language, and were treated exactly as established English-origin loanwords in Arabic, and in many occasions like native Arabic nouns, a finding that supports the nonce borrowing hypothesis.

### 2. Data, subjects, and methodology

The data which was extracted from approximately twelve hours of tape-recorded natural conversations, was collected from seven native speakers of Gulf
Arabic\(^1\) (GA), aged between 24-35. These informants share similar degrees of proficiency in English. At the time of the interviews each of the informants had been living in Canada for at least three years, and each of them had at least eight years of formal training in English in her home country. To avoid any influence on the linguistic behaviors of the informants, the participation of the interviewer was kept to a minimum.

Being an in-group member, the interviewer was able to elicit the most spontaneous data that could possibly be uttered by those informants. This type of interview-in-group- is reported to be the only possible method for obtaining the required data for such bilingual studies (Poplack, and Meechan, personal communication). This method was also used in other studies such as Eze (1998), Samar and Meechan (1998), Budzhak-Jones (1998), Poplack and Meechan (1995).

The English-origin nouns that occurred in the corpora were divided into four categories:

**i. Established English borrowings in Arabic (loanwords):** single English-origin nouns that are characterized by full integration into Arabic and by widespread diffusion even among monolinguals. Each English-origin noun that has an entry in the native dictionary Qafisheh (1996) or occurred more than four times in the corpora, uttered by two speakers or more, was grouped under this category. An example of an English loanword in Arabic is *le:san* ‘license’ in (2), which shows the phonological integration of the noun into the system of GA. The diphthong [\(\alpha y\)], which does not exist in the phonological system of this variety, is substituted by the native long vowel [e:]. Also, the second syllable of this word which is (CVCC) changes to (CVC) to conform with the phonactics of this variety which does not allow syllables of the type (CVCC) to occur at the right edge of words that are not monosyllabic (Bukshaisha 1985, p. 6).

(2) intibah-t inna il-\(le:san\)^2 mu f-il-bu:k. (10B-120D)
   realized-I that the-license not in-the-purse
   ‘I realized that the license was not in the purse’

**ii. Lone English-origin nouns:** single English-origin nouns that occurred less than five times in the corpora, have no entries in native dictionaries, and are not

---

1 This variety of Arabic is spoken in the Arab states of the Persian Gulf: Bahrain, Kuwait, Qatar, east coast of Saudi Arabia and the United Arab Emirates.
2 Beside its phonological integration, this noun is also morphologically integrated into the system of Arabic as shown by the plural form it takes: *liya:si:n*, which is one of the broken plural templates in GA.
widespread in the Gulf communities since they do not occur in the speech of monolingual speakers of the variety. The focus of the current study will be on the behavior of the nouns in this category, since these are the nouns that are considered by some scholars to be codeswitches (Bentahila and Davies 1983;1991, Bokamba 1988, Eliasson 1994, Mansour 1998, Nortier 1990, Myers-Scotton 1995, Mohamed 1989, Sallo 1994, Hussein & Shorrab1993, Atawneh 1992), and by others to be borrowings (Poplack 1993, Poplack and Meechan 1995;1998, Poplack, Wheeler and Westwood 1989, Budzhak-Jones 1998, Samar and Meechan 1998, Eze 1998, Turpin 1998), the identity of these nouns are kept ambiguous until their pattern of behavior show similarities to either established loanwords in Arabic, or unambiguous codeswitches. Examples of this category of nouns are *profisor* and *profisOr*:t-ha in (3):

(3) kān-at ʕala ʕilaqa  maʕa profisOr min profisor-α:t-ha maɭ was-she on relationship with professor of professor-pl.her of iS- Saydala. (3A-090N)
the-pharmacology.
‘she was having an affair with one of her pharmacology professors’

**iii. Unambiguous codeswitches**: these are English nouns that occurred in multi-word fragment of English which are bordered from at least one side by Arabic items.

(4) w ihna:k  yaʕT-u:ni-ch your license. (2A-473M)
and there give-they-you
‘and there, they give you your license’

**iv. Monolingual English**: English nouns that occurred in larger constituents of English are considered to have occurred in monolingual English.

(5) living with a chronic illness, basically living with something chronic. (1B-206M)

**Monolingual Arabic**: Arabic nouns that occurred in constituents of Arabic are considered to have occurred in monolingual Arabic.
From two of the interviews, a number of Arabic nouns in monolingual Arabic contexts were also extracted for comparison purposes.
(6) *Tila* il-yo:m mihabbib. *(1A-404M)*

appeared the-day awfu
‘the day appeared to be awful’

The comparative methodology that was used in this study is based on the idea that the grammatical (not etymological) membership of other language items can be known by comparing the patterns of behavior of these items to that of their counterparts in the monolingual context of each language involved, to their counterparts in unambiguous codeswitches, and to established borrowings (loanwords). So, to reveal the identity of lone English-origin nouns in Arabic contexts, these nouns were compared to Arabic nouns in monolingual Arabic, to established English loanwords in Arabic, to English nouns in multi-word fragments of English (unambiguous CSs), and to English nouns in monolingual English, see figure (1). The total number of nouns in each category is given in Table (1).

Figure (1) The comparative methodology.

<table>
<thead>
<tr>
<th>Arabic nouns in monolingual Arabic.</th>
<th>English nouns in monolingual English.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lone English-origin nouns in Arabic contexts</td>
<td></td>
</tr>
<tr>
<td>Established English loanwords in Arabic.</td>
<td>English nouns in unambiguous Codeswitches to English.</td>
</tr>
</tbody>
</table>

Table 1 The total number of nouns in each category.

<table>
<thead>
<tr>
<th>Category</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic nouns in Arabic contexts</td>
<td>287</td>
<td>22</td>
</tr>
<tr>
<td>Established borrowings (loanwords)</td>
<td>526</td>
<td>41</td>
</tr>
<tr>
<td>Lone-English-origin nouns</td>
<td>226</td>
<td>18</td>
</tr>
<tr>
<td>Codeswitched nouns</td>
<td>162</td>
<td>13</td>
</tr>
<tr>
<td>English nouns in English contexts</td>
<td>86</td>
<td>7</td>
</tr>
<tr>
<td>Total</td>
<td>1287</td>
<td></td>
</tr>
</tbody>
</table>
To apply the comparative methodology, conflict sites between the two languages were investigated. Conflict sites are sites at which the structures of the two languages do not match (Poplack and Meechan 1998, p. 133). These sites are supposed to show the language membership of lone words. That is, when lone English-origin nouns follow the structure of one of the two languages in these sites, it follows that these lone words are following the grammar of that language, hence, are members of this particular language. The conflict sites that I investigated are patterns of determination, gender assignment, and word order.

The nonce borrowing hypothesis predicts that if lone English-origin nouns in otherwise Arabic contexts follow the grammar of Arabic and pattern like native Arabic nouns as well as the established loanwords in Arabic, while at the same time differing from the English nouns whether in unambiguous codeswitches to English, or in monolingual English discourse, these lone items are behaving like borrowings not codeswitches. If, on the other hand, these lone items follow the grammar of English, patterning with English nouns in English discourse as well as English nouns in unambiguous codeswitches to English, then they are functioning as codeswitches.

Using Arabic/English bilingual discourse, and applying the comparative variationist methodology developed by Poplack and Meechan (1995) for this purpose, I verify the predictions of the nonce borrowing hypothesis with respect to single English-origin nouns in Arabic contexts.

3. Conflict sites

3.1. Determination

In Arabic, the definite determiner occurs in more contexts than that of English. One of these contexts is the generic context. Unlike English, in Arabic, these nouns must be definite or added to other definite nouns. A generic use of the word ‘mara’ meaning woman is in (7), where il- the definite determiner precedes ‘mara’:

(7) o ha-sh-shay mu marghoub ?ind il-mara. (7A-009B)
   and this-the-thing not wanted near the-woman
   ‘and women wouldn’t like that’

Also, Unlike English, in this variety of Arabic, indefiniteness is normally indicated by zero, which means, the determiner is absent, as shown in (8), where filim is marked by zero.
Each noun was coded for whether it was modified by a definite determiner, as ‘language’ in (9):

(9) bas *il*-language muhim. (1A-645M)
   but the important.
   ‘but language is important’

or an indefinite determiner as *mara* in (10), which is preceded by *waHda* ‘one’.

(10) cha:n tiTla’i la-ha *waHda* mara ikbira. (7B-070N)
   ‘then she finds an old woman’

If the patterns of determination of the single English-origin nouns show similarity to those of native Arabic nouns and established loanwords in Arabic, the nonce borrowing hypothesis predicts that these lone items have been borrowed, even though momentarily. If, on the other hand, English-origin nouns show similarities to English nouns in monolingual English as well as in unambiguous codeswitches to English, then we conclude that they are behaving like codeswitches.

Table (2) The patterns of determination of the nouns in each category.

<table>
<thead>
<tr>
<th></th>
<th>Arabic nouns</th>
<th>Established loanwords</th>
<th>Lone items</th>
<th>English nouns in codeswitches</th>
<th>English nouns in English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Definite determiner</td>
<td>121</td>
<td>97</td>
<td>237</td>
<td>99</td>
<td>93</td>
</tr>
<tr>
<td>Indefinite determiner</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>125</td>
<td>95</td>
<td>239</td>
<td>95</td>
<td>45</td>
</tr>
</tbody>
</table>

Table (2) shows that English-origin nouns were modified by the definite determiner in a pattern that is almost identical to native Arabic nouns and established loanwords (98%, 97% and 99% respectively). Moreover, these lone items were treated differently from their counterparts in codeswitches and in
monolingual English ((78% and 39%). The same thing is true of the indefinite determiner. This table shows that the lone items were treated similarly to native Arabic nouns and established loanwords (2%, 3%, and 1% respectively), and that they were treated differently from their counterparts in codeswitches and in monolingual English (22% and 61% respectively).

The behavior of the lone English-origin nouns in determination contexts suggests that these nouns are borrowed into the system of Arabic which is the prediction made by the nonce borrowing hypothesis.

3.2. Gender

In English, the majority of inanimate nouns are neuter, but in Arabic, animate and inanimate nouns are assigned either feminine or masculine gender, depending on the final phonetic sound of the noun. This strategy for gender assignment is used with English-origin nouns as well.

In Arabic, the gender of a given noun is indicated by the surrounding items such as pronouns, clitics, verb inflection, adjectives, and demonstratives. If the context does not indicate the gender that is assigned by the speaker to a particular noun in a particular sentence, the noun is considered not to be assigned gender, an example is ‘rashwa’ in (11):

(11) gal-o-l-a idfaʕ la-h rashwa. (1A-368B)
    told-they-to-him pay to-him briber
    ‘they told him: bribe him’

But ‘juice’ in (12) was assigned a masculine gender, as indicated by the verb inflection:

(12) liʔana il-juice aSlan ʕumri ma HaTeit-ʕa f mirthaʕa. (10-A-058L)
    because the- really never no put-it in bottle
    ‘because I really never put juice in a bottle’

According to the nonce borrowing hypothesis, if the pattern of gender assignment to the lone English-origin nouns follows that of native Arabic nouns, and established loanwords in Arabic, these lone items would be behaving like Arabic, hence, must be considered to be borrowed into the system of Arabic. If, on the other hand, they follow the pattern of gender assignment to English nouns in monolingual English, and in codeswitches to English, the lone items must have
been codeswitched. The results of the analysis of gender assignment are given in table (3).

Table (3) The patterns of gender assignment to the nouns in each category

<table>
<thead>
<tr>
<th>Gender indicated</th>
<th>Arabic nouns</th>
<th>Established loanwords</th>
<th>Lone items</th>
<th>English nouns in codeswitches</th>
<th>English nouns in English</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
</tr>
<tr>
<td>Gender indicated</td>
<td>42</td>
<td>19</td>
<td>85</td>
<td>20</td>
<td>49</td>
</tr>
<tr>
<td>Gender not indicated</td>
<td>184</td>
<td>81</td>
<td>347</td>
<td>80</td>
<td>154</td>
</tr>
<tr>
<td>Total</td>
<td>226</td>
<td>432</td>
<td>203</td>
<td>141</td>
<td>76</td>
</tr>
</tbody>
</table>

Table (3) shows that the gender of the noun was indicated for lone nouns, established loanwords, and native Arabic nouns, at similar rates (24%, 20%, and 19%). It also shows that 76% of the lone nouns were not assigned Gender. Since Arabic is a language that assigns gender to all kinds of nouns, this may indicate that the lone nouns were governed by the grammar of English not Arabic, hence, they were codeswitched. But the fact that established loanwords and even native Arabic nouns were not assigned gender at even higher rates (80% and 81% respectively) can easily eliminate this possibility, especially since English nouns in codeswitches were almost categorically not assigned Arabic gender. This indicates that codeswitches were following the grammar of English. On the other hand, lone items were subject to the same strategies in gender assignment as the established loanwords and native Arabic nouns, which is an indication that these nouns were incorporated into the system of Arabic and following the grammar of Arabic. And since it is obvious that there is a difference in the treatment of established loanwords and lone items on the one hand, and unambiguous codeswitches on the other, they must be undergoing different processes.

3.3. Word order

English is an SVO language. On the other hand, the word order in Arabic can be SVO, VSO, or even VOS in certain cases. An example of a VSO word order is (13), where the verb ‘yinzil’ precedes the subject ‘snow’:

(13) ga: ?id yi-nzil snow yazir. (3A-217B)
    progressive it-falls snow heavy
    ‘it is snowing heavily’
(13) shows another disagreement between the two languages, that is, adjective/noun placement. Following the grammar of Arabic, the adjective ‘yazir’ follows the noun.

Also, the two languages conflict in possessive noun placement. In English, possessive pronouns precede the noun, whereas in Arabic they follow it. This is illustrated in (14), where ‘maːli’ follows vacation:

(14) a-xalli il-vacation maːli Hag next year. (7B-333L)
I-keep the- of-me for
‘I’m saving my vacation for next year’

If the word order of the structures that the lone items occurred in shows a pattern similar to that of the nouns in unambiguous codeswitches and monolingual English, as opposed to the patterns of the structures that the established loanwords and native Arabic nouns occurred in, these lone items would be best characterized as codeswitches. If, on the other hand, the lone items follow a pattern that is similar to native Arabic nouns and established loanwords, the nonce borrowing hypothesis would characterize them as borrowings.

The results of word order are given in table (4).

Table (4) the patterns of word order of the nouns in each category.

<table>
<thead>
<tr>
<th>Arabic nouns in Arabic</th>
<th>Established Loanwords</th>
<th>Lone items</th>
<th>English nouns in codeswitches</th>
<th>English nouns in English</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>%</td>
<td>N</td>
<td>%</td>
<td>N</td>
</tr>
<tr>
<td>Follow Arabic</td>
<td>85</td>
<td>30</td>
<td>139</td>
<td>27</td>
</tr>
<tr>
<td>Follow English</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.2</td>
</tr>
<tr>
<td>Follow both</td>
<td>202</td>
<td>70</td>
<td>386</td>
<td>73</td>
</tr>
<tr>
<td>Total</td>
<td>287</td>
<td>526</td>
<td>526</td>
<td>73</td>
</tr>
</tbody>
</table>

The lone items followed word orders that are specific to Arabic 29% of the time, which is almost identical to the behavior of both native Arabic nouns, 30%, and established loanwords, 27%. None of the unambiguously codeswitched nouns followed these word orders. On the contrary, the nouns in unambiguous codeswitches occurred 70% of the time in structures that followed the word order of only English, which is even higher than the rate of the English nouns in monolingual English contexts 62%.
Even when the word order complied with both languages, the lone items behaved identically to native Arabic nouns and very similarly to established loanwords, but differently from the nouns in unambiguous codeswitches and monolingual English. This is further proof that each of borrowing and codeswitching is a different process and that the lone items were treated as borrowings, not codeswitches.

4. Conclusion

By investigating the patterns of determination, gender assignment, and word order of English-origin lone items in otherwise Arabic discourse, I was able to show that these lone items, as well as the established English loanwords in Arabic are all the “outcome of the same mechanism”: The patterns of behavior of lone items were not only similar to that of established loanwords in Arabic and native Arabic nouns, but they were also different from the patterns of behavior of unambiguous codeswitches, which indicates that lone items must be characterized as borrowings, not codeswitches, and that borrowing and codeswitches are products of different processes. It is also clear that for an item to be borrowed, it need not be widespread or recurrent.

The findings of this study support the findings of a growing body of variationist analyses of the status of English-origin items in such typologically distinct language-pairs as Tamil (Sankoff, Poplack and Vanniarajan, 1990), French (Turpin, 1998), Persian (Ghafar-Samar and Meechan, 1998), Finnish (Poplack, Wheeler and Westwood, 1989), Ukranian (Budzhak-Jones, 1998), Igbo (Eze, 1998), in contrast to English and Wolof and Fongbe in contrast to French (Poplack and Meechan, 1995). All of them show that borrowing and codeswitching are different processes and that most lone items are the outcome of borrowing not codeswitching.

References


Actes de l’ACL 2002/ 2002 CLA Proceedings
Poplack, Shana, Sankoff, David, & Miller, Christopher. (1988). ‘The social correlates and
linguistic processes of lexical borrowing and assimilation’. Linguistics, 26, 47-104.