# Code-Switching in the Speech of Native Speakers of Arabic in the United States: Structure and Motivations

**Khaled Abu-Abbas[1], Muhammad Badarneh[1]** and **Wael Zuraiq[2]**

## Abstract

The present study investigates the phenomenon of code switching in the speech of native speakers of Arabic living in the United States. Data is analyzed within the framework of the Matrix-Language Frame Model (MLFM). The speech of two groups with distinct educational levels was recorded in informal situations. The purpose of this study is threefold: (i) considering the syntactic categories of the switches for typological comparison; (ii) the motivations behind code switching are then investigated in light of (i); and (iii) explaining the morphological composite of these switches with reference to the MLFM.

## 1. Introduction

### 1.1 Overview

The ability to speak two languages i.e., bilingualism is increasingly becoming the norm rather than the exception in any modern society. A bilingual often shifts form one language to the other almost unconsciously in a process referred to as code-switching or code-mixing (Blom & Gumperz, 1972; Castells et al., 2006; Grosjean, 1982; Myers-Scotton, 1990, 1993, 1995, 1998, 2001; Poplack, 1988; Sue, 2003 to name but a few). Bilinguals often switch back and forth between languages in the same utterance. This process takes place frequently and almost unconsciously within a single social event (Sridhar, 1978). In any speech event where the interlocutors are bilingual, code switching is the norm rather than the exception. The two language systems of the bilingual are active simultaneously (Joshi, 1985).

Most scholars use the terms Code Switching (CS) and Code-Mixing (CM) interchangeably while others distinguish between them (Auer, 1999; Kachru, 1978; 1983). A related term as well is that of Borrowing. Code-switching involves the use of two languages in one utterance, whereas the term 'borrowing' is used to refer to embedded elements that have been integrated into the host language (Pahta, 2004, reported in Crespo & Moskowich, 2006: 51).

1- Department of English for Applied Studies, Jordan University of Science and Technology, Irbid, Jordan.
2- Department of English, Hashemite University, Zarqa, Jordan.

Haugen (1973: 521) defines code switching as "the alternate use of two languages including everything from the introduction of a single unassimilated word up to a complete sentence or more in the context of another language". This definition was introduced in order to include single lexemes as possible material available for code switching. Earlier research limited code switching to phrases and sentences (Reyes, 1976). Single lexemes were seen as cases of borrowing that are different from code switching. In his definition of code switching, Haugen (1973) stresses the fact that single words might be subject to code switching. He basically distinguishes between two types of single words from one language used in the context of another: words that undergo the assimilation rules of the host language, and words that resist such rules. The former type constitutes cases of borrowing and the latter cases of code switching (section, 3).

In any situation of code switching, one needs to distinguish between the two types of languages involved. According to Myers-Scotton (1993), CS involves a Matrix Language (ML), which is the dominant language providing most of the morphemes, and an Embedded Language (EL), which is an intruder providing linguistic material to the conversation. A detailed discussion of the two language types and their contribution to the code switching event is presented in section (7) below.

The term 'bilingual' will be used loosely here to refer to people who speak two languages with some proficiency. Myers-Scotton (1993) notes that speakers need not be entirely fluent in the EL when they engage in CS. Rather, they need to know some content morphemes from the EL in order to use them in CS, but they do not need to master the system morphemes or phrase structure rules that are necessary to produce well-formed sentences in the EL. They need to be proficient in the EL only if they are to produce well-formed phrases and sentences in the CS situation.

In this paper, we will follow Myers-Scotton's definition of CS who states (1993: 3) that "code-switching is the selection by bilinguals or multilinguals of forms from an embedded variety (or varieties) in utterances of a matrix variety during the same conversation".

## 2. Intersentential vs. intrasentential CS

Intersentential CS refers to instances of CS between sentences. A speaker would suddenly switch to the EL after a whole sentence was produced in the ML. Intrasentential CS on the other hand refers to instances of CS within the boundaries of a single sentence. The MLFM is primarily concerned with intrasentential rather than intersentential CS.

## 3. CS vs. borrowing

Much of the research on CS excludes single lexemes from the discussion as such words are considered cases of borrowing rather than CS. Research on CS was limited to phrases and sentences (Reyes, 1976). Other research has attempted to differentiate between single words that are cases of borrowing and those that are cases of CS. As mentioned earlier, Haugen (1973) suggests that single words that undergo the assimilation rules of the ML are to be seen as cases of borrowing (this is termed 'interference' by Rouchdy (1992)) while those that resist such processes are cases of CS. Others (Poplack and Miller, 1988) still expand on the idea of assimilation, suggesting three levels of integration to differentiate between CS and borrowing. These three levels are phonological, morphological, and syntactic[1]. The notion of 'Interference' or 'Assimilation' within CS is elaborated in section (6.3). However, Myers-Scotton (1993) rejects these differentiations between CS and borrowing and argues that, within an MLFM, instances of CS and borrowing behave the same. The basic difference for her is that borrowed words tend to become part of the speaker's lexicon. They are more readily accessible and tend to be used more frequently. In this paper, we will follow the view that borrowed words are those that have no equivalent in the ML. These basically include place names, specific products, and culture-specific terms.

We believe that a further distinction between cases of CS and cases of borrowing lies in the fact that borrowing is performed by bilinguals and monolinguals alike. The borrowed words become an integral part of the lexical system of the language so much that they are rarely seen as foreign. Examples of such words in Arabic are: internet, virus, computer, CD, radio, telephone, brake, steering, gear, just to name a few. A detailed discussion of such words is beyond the scope of this study and will not be pursued any further.

## 4. Data

The data used for this paper comes from two sources. The first is a two-hour recording of informal gatherings of six male native speakers of Arabic from Jordan studying at the University of Kansas (henceforth, KU), USA. Their ages ranged from 22-35. Three of them had spent a period of four years in the States, and the other three had spent two years. Five of the six subjects are graduate students and the sixth is a senior undergraduate. The subjects were aware that they were being recorded but were not told the exact purpose of the recordings. They were asked to behave as naturally as possible and were told that the subjects discussed will not be the focus of this research. The second source is a two-hour recording of informal gatherings of Arab cab drivers in Chicago. The recordings took place at Chicago-Midway Airport, and were over a period of one week involving five male subjects ranging in age from 24 to 40. Four of them

are Jordanians and the fifth is a Palestinian. Four of the subjects had been in the States for five years and the fifth had been there for just two years. Three of the subjects had a high school and the other two finished preparatory school diploma from their native country. Once again the subjects did not know that exact purpose behind the recordings.

## 5. Data analysis

The data recorded featured a good number of instances of CS between Arabic, the ML, and English, the EL, a categorization which will be discussed when introducing the MLFM. Exchanges involving CS were extracted from the recordings and phonemically transcribed. Lexical and phrasal code switches were calculated. Instances of lexical CS were divided according to their syntactic categories, i.e., nouns, verbs, adjectives, etc. Multi-word code switches were also divided into their syntactic categories, i.e., noun phrase, verb phrase, etc. Instances of intrasentential code switching were then considered within the assumptions of the MLFM.

The rest of the paper is organized as follows. Section 6 considers the types of syntactic categories that are frequently subjected to CS. The results will be compared to the findings of other studies. Though this contribution may seem redundant since most research on the topic confirms very similar switching patterns, it is, however, crucial in this research for achieving rest of the objectives. The educational background of the participants is then related to the quantity and quality of their switches. The two groups exhibit distinct switching patterns in terms of the number of nouns vs. noun phrases on the one hand and the number of longer stretches going beyond a phrase from the EL on the other. In light of the types of words switched, section 7 considers the motivations behind code switching. Finally, in section 8 we analyze the data collected in light of the Matrix-Language Frame Model (MLFM) proposed by Myers-Scotton (1993, 1995, 1998, 2001). This will in effect explain the phenomenon of interference or assimilation which, according to Rouchdy (1992:36), "occurs when grammatical rules of the dominant language affect grammatical rules of the subordinate or borrowing language". Interference is briefly discussed in (6.3).

## 6. Syntactic categorization

Most of the research on the syntactic categorization of CS compares the frequency of single-word switches, or lexical switches, and phrasal switches (Lindholm & Padilai, 1978; Sridhar and Sridhar, 1980; Poplack, 1980; Houwer, 1990; Eid, 1991; Atawneh, 1992; Bader and Minnis, 2000). Another body of literature on CS has been structurally oriented. It attempts to answer the question: Where in a sentence might a speaker switch from one linguistic variety

to another? Syntacticians looked for the surface structure of sentences for clues about the constraints on CS like word order, surface category membership, and size of the constituent being switched (Gumpers, 1982; Heller, 1988). As for the syntactic categorization of the switches, nouns were found to be cross linguistically more often switched compared to any other lexical category and so were noun phrases at the phrasal level. The other categories seem to vary among languages.

## 6.1 Lexical vs. phrasal code switching

The percentages provided are calculated out of the total number of switches, i.e., lexical and phrasal. Table 1 provides the number of lexical and phrasal CS obtained from the speech of the Chicago cab drivers.

Table (1) shows that single-word switches (77.17%) are more frequent than phrasal switches (22.83%). The most frequent switch among single words is the noun (85.91%), whereas among the phrasal category it is only the noun phrase (100%). Verbs were never switched, adjectives were infrequently switched (11.27%), and only two adverbs were switched (2.82%). Almost all previous research confirms the fact that lexical categories are more frequently switched compared to phrasal categories, and that nouns and noun phrases are the highest among the two categories. For example, the frequency of occurrence, from highest to lowest, was found to be nouns, adverbs, adjectives, and verbs (Lindholm and Padilla, 1978), nouns, adjectives, adverbs and verbs (Sridhar & Sridhar, 1980), nouns, adjectives, verbs, and adverbs (Houwer, 1990), and nouns, verbs, adjectives, and adverbs (Atawneh, 1992). Accordingly, the results shown in Table 1 come in support of all previous research in that nouns are the most frequently switched category in the speech of our subjects. No verbs were switched in the two-hour recording.

**Table 1**: Percentages of lexical and phrasal CS among Arab cab drivers in Chicago.

| Lexical Categories | | | Phrasal Categories | | |
|---|---|---|---|---|---|
| Category | switches | % | Category | switches | % |
| Noun | 61 | 85.91 | NP | 21 | 100 |
| Verb | 0 | 0 | VP | 0 | |
| Adj | 8 | 11.27 | Adj.P | 0 | |
| Adv. | 2 | 2.82 | Adv.P | 0 | |
| Prep. | 0 | 0 | PP | 0 | |
| Total | 71 | 100 | Total | 21 | 100 |

Grand total of switches in both categories: 71+21= 92

Percentage of lexical switches: 71/92= 77.17%

Percentage of phrasal switches: 21/92= 22.83 %

155

It is interesting here to note that prepositions and articles were never switched. This may be accounted for by the fact that these categories belong to closed classes of words in any language while the preferred categories for switching belong to open classes of words (Joshi 1985). All phrasal categories that were switched involved the definite article in Arabic. This will be further investigated in light of the MLFM.

Table (2) below provides information on cases of code switching obtained from a two-hour recording of informal gatherings of Arab students at KU:

**Table 2**: Percentages of lexical and phrasal CS among Arab students at KU

| Lexical Categories | | | Phrasal Categories | | |
|---|---|---|---|---|---|
| **Category** | **switches** | **%** | **Category** | **switches** | **%** |
| Noun | 43 | 78.18 | NP | 23 | 100 |
| Verb | 2 | 03.64 | VP | 0 | 0 |
| Adj. | 6 | 10.91 | AdjP | 0 | 0 |
| Adv. | 4 | 07.27 | AdvP | 0 | 0 |
| Prep. | 0 | 0 | PP | 0 | 0 |
| Total | 55 | 100 | Total | 23 | 100 |

Grand total of switches in both categories: 55+23= 78

Percentage of Lexical switches: 55/78= 70.51%

Percentage of Phrasal switches: 23/78= 29.49

Table 2 confirms the results in Table 1. Lexical words are more frequently switched than phrasal categories. Nouns and noun phrases are again the highest among the two categories, with adjectives coming in second place, adverbs in third place. From the phrasal categories, once again it was only the noun phrase that underwent CS. This comes in total agreement with the results found in Table 1.

### 6.1.1 Comparing the results

A closer look at the two tables above reveals subtle variations underlying the apparent conformity in results. The total number of switches is rather close. Chicago cab drivers (Chicago group) tended to code switch a little more often compared to KU students (KU group). Both groups avoid switching prepositions in isolation and all phrases except for noun phrases. The Chicago group tends to switch lexical words (77.17%) more than the KU group (70.61). The difference might not be significant, but more variation is evident when comparing the switching of nouns. 66.30% (61/92) of the total number of switches for the Chicago group involved nouns while the KU group had 55.13% (43/78) switches involving nouns. On the other hand, the KU group switched more noun phrases 29.49 %(23/78) compared to the Chicago group who had 22.83% (21/92) of

their total switches as noun phrases. The KU group also switched more adverbs (7.27%) compared to the Chicago group (2.82%).

Below are two examples of CS[2]. In (1), a cab driver tells a story about some of the customers he met while working in a store, while in (2), a KU student tells a story about what he did with his students in a computer programming course he was teaching:

(1) bidži-    ik      zbuun    biwaggif    ʕal. <u>lajn</u>    saaʕa   bidd-o
    comes/    you     customer stands      on the /line      hour    want-he

    <u>čeindž for ə dalər</u>    waraah        waħad              θaani   maʕaa   ʔrbaʕ
    change for a dollar  behind/him    one                second with/him four

    <u>qwaara-at</u>        w-bidd-              o <u>solid dalər.</u>
    quarter-plural       and/want/he          solid dollar

    'Sometimes a customer comes in and stands in line for an hour wanting change for a dollar. Another right behind him has four quarters and wants a solid dollar.'

(2) daxal-t    ʕale-hom        il.joom  miš   mħaDDer         w- ʔolt-il-hom
    Entered-I  on- them        the-day  not   prepared        and/said/to/them

    ʔismaʕ- u          baħoT-k- u     fi    <u>groups</u>         w-ʕale- k- u
    Listen-plural      put- you-plural in    groups           and/have/you-plural

    tsaw-u      <u>kəmpjuutər pleiərz</u>   la-liʕbe   w- bidd- ii- š   <u>raandəm pleiərz</u>
    do-plural   computer players       for-game   and/want/I/not  random players

    jkuun-u     <u>stupid</u> bidd-i  <u>inteLədžənt</u>          <u>kəmpjuutə</u>        <u>pleiərz</u>
    be/ them    stupid   want/I  intelligent           computer        players.

    'I went to them today not prepared and said to them: "Listen, I'll put you in groups and have you do computer players for the game, and I do not want stupid random players, I want intelligent computer players.'

## 6.2 Sentential CS

As noted in the literature, CS might take place at the lexical level, phrasal level, or beyond to include a whole sentence or even several sentences. The more elaborate the switched item, the more proficiency the speaker is required to have in the EL. Proficiency might be achieved through long years of exposure to the language. The data collected for this study featured subjects who roughly had the same period of exposure to English as a second language. Nevertheless, the two groups seem to have variations when it comes to switching a whole sentence

or several sentences. This might suggest that there are factors other than exposure time that influence a speaker's ability to switch structures larger than a single word. Another note that supports this claim has to do with the quality or the structural complexity of the switched sentence or sentences.

In the two-hour recording of casual speech among Chicago cab drivers, there were four instances of switching that went beyond a single word or phrase, two of which involved reporting conversations with native speakers of English and the other two were about a football game the drivers were watching. Two of these switches are presented in (3) and (4) below:

(3) waggaf-it   ʕind-ha     w – gul-it      <u>Hi, how are you? Do you need a cab</u>?

     stopped/I   near/her     and-said

     gaal-at we don't need no cab.

     said/she

(4) (Talking about a football player): He averages five yards <u>bil</u>(in the) carry.

On the other hand, during the two-hour recording of casual gatherings among KU students, there were thirteen instances of switches that involved a sentence or more. Some of these switches are shown in (5-10) all of which occurred within conversations in Arabic:

(5) They are very active politically.

(6) Just give me a few seconds to think.

(7) You can take it up to four times a day.

(8) I produce the image; a photo realistic image.

(9) I'm just kidding.

(10) The more you ask them (the students) to do, the less you do. That's actually a good way to learn programming. To have them sit down and do it. It's like mathematics. You have to let them sit down and do it.

A comparison of the quality and quantity of the switches performed by the two groups shows that the Chicago group switches less and tends to incorporate ML morphemes to break the sentences. The KU group clearly opts for sentential code switching more often. The examples listed from the KU group show that none of the examples were interrupted by Arabic morphemes or words. Example (10) above shows the level of complexity of some multi-sentence CS.

## 6.3 Interference

Interference occurs when grammatical rules of the dominant language affect grammatical rules of the subordinate language (Rouchdy, 1992). In the case considered in this study, phonological rules of Arabic are applied to English lexical items. Cases of interference will be more evident as the differences between the two languages involved increase. The non-concatenative nature of Arabic morphology (McCarthy, 1979) implies that there should be many cases of interference when native speakers of Arabic switch to English, which is, for the most part, a language with a concatenative morphology. For example, the analysis shows that native speakers of Arabic tend to use the Arabic feminine sound plural marker [-aat] i.e., the regular plural marker, rather frequently with English singular nouns as shown in (11).

(11) The Arabic regular plural in CS

| | | |
|---|---|---|
| a. kaabaat | kaab pl. | cabs |
| b. džakkaat | džak pl | jacks |
| c. trakkaat | trak pl | trucks |
| d. billaat | bil pl | bills |
| e. storaat | stor pl | stores |
| f. keekaat | keek pl | cakes |

One can claim that this morpheme is very productive in CS. There are other cases where the speaker would opt for certain broken plural forms, i.e., an irregular plural marker, rather than the unmarked feminine sound plural. This would be an interesting research topic since there are 32 different shapes, i.e., templates, for the broken plural in Arabic.

The Arabic definite article was also used in most cases instead of the English counterpart. The assimilation rule associated with the Arabic definite article where the definite article /l/ fully assimilates with a following coronal sound is also active in examples involving English nouns as shown in (12).

(12) Assimilation of the Arabic definite article

a. /ʔissors/ 'the source'    b. /ʔiljuzər/ 'the user'    c. /ʔissiin/ 'the scene'

d. /ʔiššado/ 'the shadow'    e. /ʔilmaws/ 'the mouse' f. /ʔillajn/ 'the line'

The use of the definite article with English nouns was found to be very common and can be taken as a productive process. For every instance of a definite article within a phrase, the Arabic definite article was used. The English definite article was only preserved within sentential CS. The assimilation rule was also very productive in the data, i.e., one would be able to identify the

English nouns that would involve an assimilation of the Arabic definite article. For example, if a native speaker of Arabic wants to use an English noun phrase like 'the night' with the Arabic definite article, we can expect the assimilation process to take effect since /n/ is a coronal sound. According to the assimilation rule in Arabic, the /l/ of the definite article /ʔal/ assimilates to a following coronal sound.

## 7. Motivations behind CS

Although the process of CS is considered to be the norm rather than the exception when the interlocutors speak more than one language, not all cases of CS have logical motivations behind their use. This is evidence for the claim that code switching is a process that is performed unconsciously (Sridhar, 1978). Accordingly, a discussion of the motivations underlying the process of CS is of merit. We will be concerned here with two basic types of CS, which are (i) technical CS and (ii) socio-cultural CS. Within each type, some examples of CS can be justified and others seem to have no particular motivation other than the unconscious nature of the process itself.

Previous research (e.g., Castells et al., 2006; Bautista, 1999; 2004; Sue, 2003; Scotton, 1993) has named common factors that affect an individual's reasoning for code-switching. Among these are: 1) environmental settings, 2) audience, 3) conversations with embarrassing or uncomfortable topics, and 4) picking up "cues" from others that serve as an invitation to speak both languages (reported in Benitze, 2008). Functional studies of CS have focused on the reasons why bilingual speakers engage in CS, in describing the social context in which such utterances are produced, the social functions they attempt to serve, and the sociolinguistic factors triggering this kind of behavior (e.g., Chung, 2006; Castells et al., 2006; Sue, 2003; Al-Khatib & Farghal, 1999; Auer, 1999; Adendorff, 1996; Myers-Scotton, 1995; 1998; Mustafa & Al-Khatib, 1994; Gumperz & Hernandez-Chaves, 1978).

## 7.1 Technical CS

A technical term here is one that either lacks a corresponding term in the host language, or the corresponding term is rarely used and would sound odd and deprive the original English term of shades of its meaning. Examples from conversations between KU students are provided in (1) where only the terms in question are sited:

(1) Technical CS by KU students

    a- Internet    b- Graphics    c- Photo realistic image    d- Laptop   e- Hard disk

Instances of technical CS used by the Chicago cab drivers were basically related to work as shown in (2):

(2) Technical CS by the Chicago cab drivers

a- Short trip   b- Shared ride   c- Starter   d- Meter and a half.

The above examples are considered technical because they have specific implications to cab drivers. For example, a short trip is defined by the state in terms of location and time needed by the cab driver to be back at the airport. A certain trip might be closer than another, but still the latter is considered a short trip because it should take the cab driver less time to get to the destination. Further technical implications involved in the term 'short trip' is that the driver who gets the short trip and is back at the airport within the designated time does not have to stay in line to get another trip. Also a term like 'meter and a half' is considered technical because it refers to specific locations in the city of Chicago and its suburbs that entitle the cab driver to add 50% to the meter.

## 7.2 Socio-Cultural CS

Long years of exposure to a certain culture are bound to affect one's native language. Such a situation would certainly be most evident in the speech of children who learned Arabic from their parents but who have never been exposed to their native culture. Their speech will be especially marked for instances of CS induced by culture. When it comes to the speech of native speakers of Arabic who came to the United States as adults, their language would be less marked for socio-cultural CS. Nevertheless, their speech will include instances of CS where culture and the society they live in are the primary reasons for these switches. Socio-cultural CS can be divided into the following subcategories:

## 1. Place names

Although some place names can be translated since they are not names of people, still such names are always preserved in their original language. Following are some examples. The first two were said by KU students and the other four by Chicago cab drivers /Mr. Goodcents, WalMart., 555 West Adams, Down Town, Soldiers Field, Midway/.

Some of the examples above can be translated into Arabic, but such translations would be opaque even for native speakers of Arabic. For example, 'soldiers field' can be translated as [saaħat al džunuud] literally meaning 'field (of) the soldiers'. A less striking translation of a place name would be that for 'downtown'. An equivalent term is used in Arabic, which is [waSat il balad] meaning 'middle of the town'. The use of this term in Arabic would only be

taken to refer to the main street or area in an Arab country. The addressee will never take it to refer to downtown Chicago.

## 2. Other Cultural Terms

Names of products are another category that resists translation into the host language. Again the reason being that such a literal rendering would deprive the term or item of shades of its meaning that are necessary for a complete comprehension as shown in (3).

(3) Names of products

    a- Donuts  b- Bakery  c- Gap, Polo, Marshal field. (brand names for clothes)

    d- Foot long (a sandwich)   e- Italian beef (a sandwich)

    f- Mouse (computer terminology)    g- Flu shot

Once again, it would be reasonable to translate some of these terms into Arabic, but the translations will not capture the intended meaning. For example a word like 'bakery' has an equivalent in Arabic, which is [maxbaz]. The problem is that an Arab would not expect to find 'donuts' in a [maxbaz] where only bread is made. Also the compound word 'foot long' can be translated into Arabic, but it will no longer refer to that specific type of sandwich that is a foot long.

Names of places and products are not the only socio-cultural cases of motivated CS. There are other nouns that often undergo CS due to socio-cultural factors. Such nouns undergo CS simply because the equivalent term in Arabic would not capture the intended meaning. Some examples are provided in (4).

(4) Other nouns

    a- NBA (National Basketball Association)

    b- Baseball (A sport)

    c- Mileage (The number of miles a car has traveled)

    d- Short trip (For cab drivers)

    e- Super Bowl (The final game in the NFL season)

    f- NFL (National Football League)

    g- Touch down (Scoring in an NFL game)

    h- Starter (Airport employee regulating the movement of cabs)

    i- National League (In baseball)

    j- American League (In baseball)

Some of the terms above do not have formal equivalents in Arabic. American football is not played and thus the terminology associated with it does not exist. One can look for a functionally equivalent term for some of the

examples above, but still this would not be possible since the entire idea of NBA or NFL is not part of the Arabic culture. The final possible Arabic rendering of such terms would be through opting for an ideationally equivalent term or a paraphrase. This would not be a reasonable solution because no matter how good that equivalent is, it will still fail to capture the whole intended meaning, let alone the fact that a paraphrase would naturally be longer than the original term and thus CS makes sense.

## 7.3 Unmotivated CS

In spite of all the arguments concerning motivated cases of CS, there are still many instances of CS that seem to lack any reasonable justifications. For example, the use of nouns and noun phrases in English when comparable counterparts exist in Arabic that would deliver exactly the same message without confusing the audience or having to paraphrase the idea as shown in (5):

(5) Unmotivated CS

| English | Arabic |
| --- | --- |
| a- ʔil- game | ʔil-mobaaraah |
| b- ʔil- bill | ʔil-faatuurah |
| c- ʔis-supervisor | ʔil-mošrif |
| d- ʔiθ-thesis | ʔir-risaaleh |
| e- sad | ħaziin |
| f- my neck | ragbati |

The examples above would never be used when the speaker is in an Arab country even if his audience speaks English. Words like 'game', 'bill', 'supervisor', etc. have perfect counterparts in Arabic as seen on the right column. The examples of motivated CS mentioned earlier might be used when the speaker is in his home country, but they would almost always be followed by an explanation of what these terms mean.

Another interesting area that involved CS had to do with repetition. One of the subjects would start with an English word or phrase immediately followed by a translation of the term. In other cases, the speaker would start with the Arabic term followed by an English switch. This might be seen as a sort of unmotivated switch especially when the Arabic term precedes the English term. Addressing native speakers of Arabic, one cannot justify why a translation in English should follow the Arabic sentence or word. Examples are provided in (6):

(6) CS and Repetition

| a- sik.kii.net | | mat.bax | kičin najf |
|---|---|---|---|
| knife | | kitchen | kitchen knife |
| b- Sa.ħib | | ʔil.maħal… | ʔil.onər |
| owner | | the store | the owner |
| c- baštaɣil | | mo.diir | ma.nidžər |
| work (I) | | manager | manager |
| d- bjilbas | | sitizin | madani |
| wears | | (he) citizen | citizen |
| e- əm | džəst | kidiŋ | bamzaħ |
| am | just | kidding | kidding (I) |

## 8. Matrix-Language Frame Model

As proposed by Myers-Scotton (1993), this model attempts to answer the following question: When do speakers alternate between two linguistic varieties, how free is this alternation from a structural point of view? (Myers-Scotton, 1993: 1).

### 8.1. Matrix language vs. embedded language

A basic assumption of the MLFM is the fact that the two or more languages involved in the code switching situation participate differently. The language that sets the grammatical frame for any mixed constituent is called the Matrix Language (ML); the other language is called the Embedded Language (EL). In any mixed constituent, the order of the morphemes follows from the ML. This fact is referred to as the *Morpheme Order Principle*. Also, the ML provides the system morphemes in the mixed constituents. This is referred to as the *System Morpheme Principle*. According to Myers-Scotton (1995: 237), three criteria are used to determine the ML:

(i) The ML is the one that is more unmarked. Often it is the language most associated with solidarity-building functions for the speaker.

(ii) Speaker judgments point toward the ML, i.e., persons engaged in code switching can identify which language is the ML.

(iii) The relative frequency of morphemes from the participating languages is also a good indicator. The ML is the one providing relatively more morphemes.

164

These three criteria all indicate that in the data under investigation, Arabic is the ML and English is the EL. The choice of Arabic among native speakers of the language is less marked and most morphemes are taken from Arabic. During the four hours of recordings between the two groups in this study, only a handful of expressions come from English. The rest of the conversations were conducted in Arabic, clearly indicating that Arabic is the ML.

In any CS situation, there are three types of constituents that may appear in a conversation. The first is an ML constituent which is a constituent made up entirely of morphemes from the ML. In our data, most constituents belong to this category. Such constituents will not be dealt with here since such a discussion would require an extensive analysis of well-formed structures in Arabic. The second type of constituents is one that is made up of morphemes from both the ML and EL (ML+EL constituents). Such constituents need to follow the *Morpheme Order Principle* and the *System Morpheme Principle* introduced earlier and discussed in 7.2 below. Finally the third type of constituents in CS is one that is made up entirely of morphemes from the EL (EL Islands). This type is discussed in 8.4.

## 8.2 System morphemes vs. content morphemes

The basic distinction between system and content morphemes is realized through thematic role assignment. Content morphemes are either thematic role assigners or thematic role receivers. System morphemes on the other hand cannot assign or receive a thematic role.[3] System morphemes have a general deictic function, i.e., they point directly to entities or events. Any lexical item that belongs to a syntactic category which involves quantification across variables is a system morpheme (Myers-Scotton, 1995: 237). Accordingly, quantifiers, determiners, possessive adjectives, inflectional categories (person, case, and gender), and adverbs of intensity are all system morphemes that must be provided by the ML. The copula, *do*-verbs, and dummy pronouns (*it* and *there*) are all system morphemes. On the other hand, content morphemes include most verbs, some prepositions, nouns, most pronouns, and descriptive adjectives.

The ML hypothesis sketched above is responsible for determining the structural shape of constituents that may be subject to code switching. It is responsible for limiting the shape of constituents in the code switching event that are taken from the ML and EL, i.e., it controls the structure of ML+EL constituents as shown in the following case:

(1) ruħ-it        ʕa- l- beekari    w- Talab- it    donəts
    went-I        to-the-bakery    and-ordered-I  donuts

'I went to the bakery and ordered donuts'

What is of interest here is the phrase /ʕa-l-beekari/ 'to the bakery'. The word 'bakery' is a content morpheme in the EL and thus may be switched according to the ML hypothesis. On the other hand, the preposition /ʕa/ and the definite article /l/ are system morphemes that must be provided by the ML and are thus unswitched (*System Morpheme Principle*). The order of morphemes is that of the ML where the first person pronoun /-it/ is a suffix attached to the verb rather than a free morpheme preceding the verb as in English (*Morpheme Order Principle*), as shown in (2) below:

(2) haað       li- skriin      tabaʕ-tu      law     inha    ha- l - gad

    this        the-screen    own- it       if       it       this-the-size

'This (TV) screen, if it were just this big'.

Again, the lexical item /skriin/ underwent switching since it is a content morpheme in the EL. The demonstrative /haað/ resisted switching since it is a system morpheme that must be provided by the ML.

## 8.3. ML blocking hypothesis

The MLFM states that content morphemes in ML+EL constituents may come either from the ML or the EL. However, the dominance of the ML is evident even in such mixed constituents since the ML provides all the system morphemes and most of the content morphemes as well. The ML is so dominant in an ML+EL constituent that there are free morphemes in the EL that are not allowed to surface in an ML+EL constituent during code switching. This state of affairs is caused by the *ML Blocking Hypothesis*. This hypothesis limits but does not exclude EL morphemes. This is done on the basis of *congruence* between the EL morphemes and their counterparts in the ML. If a given morpheme is realized as a free morpheme in the EL, but its counterpart in the ML is a bound morpheme, the ML blocks this morpheme from appearing in an ML+EL constituent, as shown in (3):

(3) ʔiTbaʕ      ʔil-peipə      w-sajjev-ha      ʕala-disk

    Print      the-paper      and save-it      on disk

In (3), the ML provides all the system morphemes such as the definite article /ʔil/, the conjunction /w/, and the preposition /ʕala/. It also provides the content morpheme / ʔiTbaʕ/. The pronoun 'it' is a free morpheme in English. Its counterpart in Arabic is a bound morpheme [-ha]. The ML blocks the EL

morpheme from appearing in the ML+EL constituent. The shape of the verb /sajjev/ is also affected by the morphology of Arabic. This process is known as 'interference' which according to Rouchdy (1992) occurs when grammatical rules of the dominant language affect grammatical rules of the subordinate language. The process is also known as assimilation (Haugen, 1973; Poplack and Miller, 1988). In the case considered in this study, phonological rules of Arabic are applied to English lexical items. Interference will increase as the differences between the two languages increase. The non-concatenative nature of Arabic morphology implies that there should be many cases of interference when native speakers of Arabic code-switch to English, which basically has a concatenative morphology.

## 8.4. EL Islands

The ML hypothesis and the ML blocking hypothesis are sometime violated in order to produce constituents made up entirely of material from the EL. Such constituents are called EL Islands. In an EL island, all system and content morphemes are taken from the EL. An EL trigger hypothesis predicts that when a morpheme is accessed which does not comply with the dictates of the ML hypothesis and the blocking hypothesis, the linguistic material following such a morpheme must constitute an EL Island. We have seen above that the English definite article is never switched within an ML+EL constituent since it is a system morpheme that must be provided by the ML. If a speaker uses the English definite article, the entire constituent following the definite article is expected to be in English, creating an EL Island. Pronouns are also unswitched within an ML+EL constituent. Examples discussed in (6.2) show that cases of sentential CS may be viewed as EL Islands. Most examples start with pronouns or articles which are EL Island triggers as shown in (4):

(4) EL Islands

　　a. They are very active politically.

　　b. You can take it up to four times a day.

　　c. I produce the image; a photo realistic image.

　　d. The more you ask them (the students) to do, the less you do. That's actually a good way to learn programming.

Cross-linguistically, most EL Islands are quantifier phrase starting with (e.g., very, some, many), formulaic structures like (e.g., that is to say, in other words, I'm just kidding, give me a break), and adjuncts like prepositional phrases. Another major category of EL Islands found in the data is inflectional phrases (IP). This seems to be a tendency in code switching performed by native speakers of Arabic. This is due to the variation between Arabic and English in

167

their verb morphology (Myers-Scotton, 2001). Thus, the KU group switched more sentences or IPs compared to the Chicago group. In general, the KU group switched more EL islands.

## 9. Conclusion

This article considered the phenomenon of code switching performed by native speakers of Arabic living in the United States. Two distinct groups of subjects were recorded. The first group consisted of five Arab cab drivers in Chicago and the second consisted of six Arab students at the University of Kansas. The members of each group spent an average of three years in the United States.

In terms of the syntactic categories of the code-switched items, the results of the study were found to be in conformity with the previous literature in the field. Level of education of the subjects was found to play a vital role in the quality and quantity of code switching. The first group, which is less educated, tended to code-switch more nouns than the second group. They switched less noun phrases and sentences, less adverbs, and their switches were more affected by interference of the phonological rules of their native language, i.e., Arabic.

Since code switching is a natural process performed unconsciously, unmotivated cases of code switching abound. On the other hand, motivated code switching was found to involve either technical or socio-cultural terms. The morphological composite of the switches conforms to the dictates of the MLFM where the Matrix language (Arabic) provides the system morphemes and the Embedded Language (English) provides the content morphemes. Embedded Language Islands violate the rules of the MLFM since there are cases where the Embedded Language provides the system and content morphemes. The KU group used more Islands compared to the Chicago Group during code switching. Since both groups had a roughly similar exposure to English, the use of more Islands by the KU group is seen as a reflection of their higher education.

# الخلط اللغوي لدى العرب المقيمين في الولايات المتحدة الأمريكية من حيث البنية والدوافع

**خالد أبو عباس ومحمد بدارنه**، قسم اللغة الإنجليزية للدراسات التطبيقية، جامعة العلوم والتكنولوجيا الأردنية، إربـــد، الأردن.

**وائل زوريق**، قسم اللغة الإنجليزية، الجامعة الهاشمية، الزرقاء، الأردن.

## ملخص

تهدف هذه الدراسة الى الخوض في ظاهرة الخلط اللغوي لدى مجموعة من الشباب العرب في الولايات المتحدة الأمريكية حيث سيتم تحليل هذه الظاهرة في ضوء نظرية اللغة المسيطرة. تم تسجيل محادثات عفوية لمجموعتين من الشباب العرب في الولايات المتحدة بمستويات تعليمية متباينة، وعليه فإن هذه الدراسة تهدف إلى (أ) تحليل الخلط اللغوي من حيث البنية النحوية لغايات المقارنة مع الدراسات الأخرى في هذا المجال؛ (ب) تحليل الدوافع الكامنة وراء الخلط اللغوي بناء على النتائج المحصلة في (أ)؛ و(ج) توضيح البنية الصرفية للخلط اللغوي بالإشارة إلي نظرية اللغة المسيطرة.

**Notes**

1- For a detailed account, see Poplack and Miller (1988).

2- Throughout this paper, a simple phonemic transcription is used. Long vowels are written twice, /ʔ/ represents a voiceless glottal stop, and [ʕ] and [ħ] represent the voiced and voiceless pharyngeal fricatives respectively. Symbols in upper case are emphatic. Morphological details are kept to the minimum.

3- For details about thematic role assignment, see Baker (1988); Chomsky and Lasnik (1993).

# References

Al-Khatib, M. and Farghal, M. (1999). English Borrowing in Jordanian Arabic: Distribution, Functions, and Attitudes. *Grazer Linguistische Studien* 52: 1–18.

Atawneh, A. (1992). Code-Mixing in the Speech of Arabic-English Bilinguals. In Ellen Broselow, Mushira Eid, and John McCarthy (eds.), *Perspectives on Arabic Linguistics IV* , 219-241. Amsterdam/Philadelphia: John Benjamins.

Auer, P. (1999). *Code-switching in conversation: language, interaction and identity*. London: Routledge.

Bader, Y. and Minnis, D. (2000). Morphological and Syntactic Code Switching in the Speech of an Arabic-English Bilingual Child. *Multilingua* 19-4, 383-403.

Baker, M. (1988). *Incorporation: a theory of grammatical function changing*. Chicago: University of Chicago Press.

Bautista, M. (1999). An analysis of the functions of Tagalog-English code switching: Data from one case. In Maria Bautista and Grace Tan (eds.), *The Filipino bilingual: A multidisciplinary perspective (Festschrift in honor of Emy M. Pascasio),* pp. 19–31. Manila: Linguistic Society of the Philippines.

Bautista, M. (2004). Tagalog-English Code Switching as a Mode of Discourse. *Asia Pacific Education Review* 5–2: 226–233.

Benitez, S. (2008). *The Effect of Bilingualism on Text Messaging Efficiency.* Online: http://www.csudh.edu/psych/SandraBlom, Jan-Petter and John Gumperz (1972) Social meaning in linguistic structures: code-switching in Norway. In John Gumperz & Dell Hymes (eds.), *Directions in sociolinguistics*, pp. 407–434. New York: Holt, Rinehart & Winston.

Castells, Manuel, Fernandez-Ardevol, Mireia, Qiu, Jack Linchuan & Sey, Araba (2006). *Mobile Communication and Society: A Global Perspective*. Cambridge, MA; MIT Press.

Chomsky, N. and Lasnik, H. (1993). The theory of principles and parameters. In *Syntax:An international handbook of contemporary research,* ed. Joachim Jacobs,Arnim von Stechow,Wolfgang Sternefeld, andTheo Vennemann, 506-569. Berlin: Walter de Gruyter.

Chung, H.H. (2006). Code Switching as a Communicative Strategy: A Case Study of Korean-English Bilinguals. *Bilingual Research Journal* 30–2: 293–307.

Crespo, B. and Moskowich, I. (2006). Latin Forms in Vernacular Scientific Writing: Code-Switching or Borrowing? In R. W. McConchie, Olga Timofeeva, Heli Tissari & Tanja Säily (eds.), *Selected Proceedings of the 2005 Symposium on New Approaches in English Historical Lexis (HEL-LEX)*, pp. 51–59. Online: http://www.lingref.com/cpp/hel-lex/2005/paper1346.pdf.

Eid, M. (1991). Directionality in Arabic-English Code-Switching. In Aleya Rouchdy (ed.), *The Arabic Language in America*, 50-71. Detroit: Wayne State University Press.

Grosjean, F. (1982). *Life with Two Languages*. Cambridge, MA: Harvard University Press.

Gumpers, J. (1982). Conversational Code Switching. In Gumpers J. J. (ed.) *Discourse Strategies*, 59-99. Cambridge: Cambridge University Press.

Gumperz, J. and Hernandez-Chavez, E. (1978). Bilingualism, bidialectalism, and classroom interaction. In Courtney B. Cazden, Vera P. John & Dell Hymes (eds.) *Functions of Language in the Classroom,* pp. 84–108. New York: Teachers College Press.

Haugen, E. (1973). Bilingualism, Language Contact, and Immigrant Languages in the United States: A Research Report, 1956-70. In Aleya Rouch (ed.), *The Arabic language in America*, 36-49. Detroit: Wayne State University Press.

Heller, M. (1988). *Codeswitching: Anthropological and Sociolinguistic Perspectives*. Berlin: Mount de Gruyter.

Houwer, A. (1990). *The Acquisition of Two languages from Birth: A Case Study*. Cambridge: Cambridge University Press.

Joshi, A.K. (1985). Processing of Sentences with Intrasentential Code Switching. In David Dowty et al (eds.), *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*, 190-250. Cambridge: Cambridge University Press.

Kachru, B. (1978). Code-mixing as a communicative strategy in India. In James E. Alatis (ed.) *International Dimensions of Bilingual Education,* pp. 107–124. Washington, DC: Georgetown University Press.

Kachru, B. (1983). *The Indianization of English*. Oxford: Oxford University Press.

Lindholm, K. and Padilla, A.M. (1978). Child Bilingualism: Report on Language Mixing, Switching and Translation. *Linguistics* 211. 23-44.

McCarthy, J. (1979). Formal Problems in Semitic Phonology and Morphology. Ph.D dissertation, MIT, Cambridge, MA.

Mustafa, Z. and Al-Khatib, M. (1994). Code-mixing of Arabic and English in teaching science. *World Englishes* 13–2: 215–244.

Myers-Scotton, C. (1990). Codeswitching and borrowing: Interpersonal and macro- level meaning. In, Rodolfo Jacobson (ed.) *Codeswitching as a worldwide phenomenon*, pp. 85–110. New York: Peter Lang.

Myers-Scotton, C. (1993). *Duelling Languages: Grammatical Structure in Codeswitching*. Oxford: Clarendon Press.

Myers-Scotton, C. (1995). A Lexically Based Production Model of Codeswitching. In Lesley Milory and Pieter Muysken (eds.), *One Speaker, Two Languages: Cross Disciplinary Perspective on Codewsitching*, 233-256. Cambridge: Cambridge University Press.

Myers-Scotton, C. (1998). Compromise Structural Strategies in Codeswitching. In Guss Extra and Ludo Verhoeven (eds.), *Bilingualism and Migration*, 211-228. Berlin: Mount de Gruyter.

Myers-Scotton, C. and Jake, J. (2001). Explaining Aspects of Codeswitching and their Implications. In Janet Nicole (ed.), *One Mind, Two Languages: Bilingual Language Processing,* 84-116. Oxford: Blackwell Publishers Inc.

Pahta, P. (2004). Code-Switching in Medieval Medical Writing. In Irma Taavitsainen & Päivi Pahta (eds.), *Medieval and Scientific Writing in Late Medieval English*, pp. 73–99. Cambridge: Cambridge University Press.

Poplack, S. (1980). Sometimes I'll Start a Sentence in Spanish Y TERMINO EN ESPANOL: Toward a Typology of Code-Switching. *Linguistics* 18, 581-618.

Poplack, S. (1988) Contrasting patterns of codeswitching in two communities. In Monica Heller (ed.), *Codeswitching: Anthropological and Sociolinguistic Perspectives*, pp. 215–244. Berlin: Mouton de Gruyter.

Poplack, S. and Miller, C. (1988). The Social Correlates and Linguistic Processes of Lexical Borrowing and Assimilation. *Linguistics*, 26: 47-104.

Reyes, R. (1976). Language Mixing in Chicano Bilingual Speech. In J. D. Brown and J. Ornstein (eds.), *Studies in Southwest Spanish*, 182-188. Rowley, Mass.: Newbury House.

Rouchdy, A. (1992). Borrowing in Arab-American speech. In Aleya Rouchdy (ed.), *The Arabic language in America*, 36-49.Detroit: Wayne State University Press.

Sridhar, SN. (1978). On the Functions of Code-Switching in Kannda. In Braj B. Kachru & Shikaripur N. Sridhar  (eds.), *Aspects of Sociolinguistics in South Asia*, 109-117. The Hague: Mouton.

Sridhar, S.N. and Sridhar, K. (1980). The Syntax and Psycholinguistics of Bilingual Code-switching. *Canadian Journal of Psychology* 34. 407-416.

Sue, H. (2003) The Multilingual and Multi-Orthogrphic Taiwan-based Internet: Creative Uses of Writing Systems on College affiliated BBSs. *Journal of Computer Mediated Communication* 9: 1. Online: http://jcmc.indiana.edu/vol9/issue1/su.ht