Listening to English Connected Speech: A Problem and Solutions

Rajaa Aquil

Georgia Institute of Technology / School of Modern Languages

Atlanta, USA

Abstract:

Parsing connected speech has been a constant problem among learners of foreign languages. Cairene Arabic-speaking learners of English find parsing English connected speech significantly problematic. When they listen to spoken connected and running speech, they find locating and recognizing words very difficult. They cannot locate the boundaries between the words or segment the speech. This is in spite of the fact that the speech could constitute of frequent English words, which Cairene Arabic-speaking learners could know well but in citation forms. A number of phonological processes take place in connected running speech, such as contraction, consonant blending, vowel deletion, vowel epenthesis, among others. These phenomena occur based on the prosodic principles of the language, specifically syllable structure and stress. The present paper reports on a learning problem related to segmenting words from English connected speech. It reveals several pedagogical techniques that were applied to solve the problem, however, in vain. A contrastive analysis between Arabic and English is conducted and language transfer based on similarities between Arabic and English are found to be a potential force at play. The paper concludes with theoretical and applied solutions to the problem.

Key words: connected speech, stress, syllable, segmentation

1. Introduction

1.1. Connected speech

The effortlessness with which listeners recognize spoken words in a continuous speech stream contradicts the fact that it is a truly overwhelming task. Speech stream is a continuous signal in which boundaries between words may not be acoustically reliable and where the end of one word can blend into the beginning of the next.

Over the past 20 years psycholinguists have studied how L1 listeners recognize words in spoken connected speech, detect the boundaries, and know where a word starts and ends in an utterance. They found that listeners exploit various types of phonological information, such as phonological restrictions and or prosodic specifics, to segment words from continuous speech (Cutler, 1997, 1999; Cutler & Butterfield, 1992; Cutler, Dahan, & van Donselaar, 1997; Cutler & Norris, 1988; Frauenfelder & Floccia, 1999; Grosjean, 1985; Grosjean & Gee, 1987; McQueen, Norris, & Cutler, 1994; Mehler, Dommergues, Frauenfelder, & Segui, 1981; Tyler & Frauenfelder, 1987).

As for L2 connected speech, a few studies have been conducted so far. Studies conducted investigated American English -speaking learners' errors in the production of Mandarin connected speech (Chen, 2000), the processing of words by nonnative speakers of English (Hayashi, 1991), phoneme recognition and speech perception by nonnative speakers under conditions of noise (Padilla, 2003), and the perception of English phonemes in connected speech (Lenhardtova, 1993). Most of second language listening studies investigated perception of L2 phonemes. For example since the 1967s, researchers have looked into the perception of English /r/ and /l/ by Japanese speakers (Aoyamaa, Flege, Guiona, Yamada, & Yamada, 2004; Goto, 1971; Guiona, Flege, Yamada, & Pruitt, 2000; MacKain, Best, & Strange, 1981; Miyawaki et al., 1975; Mochizuki, 1981; Sheldon & Strange, 1982; Strange & Dittmann, 1984). Italian speakers' perception of English vowels has also been investigated (Flege & MacKay, 2004; Flege, MacKay, & Meador, 1999) and so has L1 Spanish speakers' distinction of

/i/ & /I/ of English and the effect of the speakers' linguistic experience on the perception of these vowels (Flege, Bohn, & Jang, 1997).

The present paper reports on a learning problem Arabic speaking, particularly Cairene Arabic speakers often have when they listen to English connected speech. The paper aims to contribute to a better understanding of the learning problem and suggest theoretical and applied solutions. It describes the learning problem and attempted pedagogical treatments. Then it provides an analysis of English and Cairene Arabic phonologies, particularly the prosodic structure of each language by highlighting the similarities and differences. Similarities in phonological aspects do not guarantee the ease of learning. As will be seen, similarities can be either facilitative or even inhibitive as discussed below. The target of this paper is to acquaint English language instructors in the Arab world of the role phonological similarities and differences have in language transfer and accordingly may affect students' learning and development of their listening skill. The paper concludes with theoretical as well as applied solutions that may contribute to the pedagogy of English specifically to teaching listening in the Arab world.

1.2. The learning problem: Difficulties for Cairene Arabic

speakers with English connected speech

Cairene Arabic (henceforth referred to as CA) speaking learners of English find connected English speech very difficult to understand and parse. During my fifteen-year experience teaching English as a Foreign Language at the American University Cairo, I found that listening had always been a strong concern of the learners. Students consistently had difficulty recognizing words that made up a connected utterance. For example, in a question like 'What do you eat for breakfast?' in natural connected speech, not every word of the utterance is stressed and fully articulated. On the contrary, a question like this is normally uttered as in (1)

1)   What do you eat for breakfast?

wʌʒʲəˈitfɚˈbrəkfæst

In English and particularly American English, which is the variety of English taught at AUC, connected utterances are shortened rather than fully articulated. Some vowels are reduced, or even deleted. Consonants are blended together, generating different allophones from the ones in slow unconnected language or in citation forms (Ladefoged, 2001). The following examples illustrate this point:

2)

a. What do you wanna eat?

wʌʒʲəwanə it

b. When'll the storm be over?

wɛnɪɫ ðə stormbiovɚ

c. You'll be fine

ɹulbiˈfain

d. Well, Italian cheese is different

wɛɫ, itæljn̩ ˈʧīzɪz ˈdɪfrənt

One of the main contributors to connected speech in English is the relationship between stressed and unstressed syllables. Stress plays a crucial role in identifying the boundaries of connected words in utterances (Cutler & Norris, 1988; Grosjean & Gee, 1987). When words generally connect or string together unstressed syllables tend to be sacrificed. Vowels in these syllables are drastically reduced, to the extent that L2 learners may not be able to identify the component words of an utterance. This inability to recognize words in English connected speech is a continual source of complaints from CA speakers who were taking English courses at AUC. Simply put, learners are not able

to match the connected forms heard in continuous English speech with their expected citation correspondents.

At AUC, I attempted various teaching techniques to address the learners' problem in listening. First, insufficient or unfamiliar vocabulary was considered to be the probable cause; hence, I made sure that students knew all the vocabulary items of the utterances they would hear. However, this did not solve the problem because neither did knowledge of English vocabulary, nor exposure to new words, seemed to help. Next, I resorted to phonology and syntax and explained to the learners the syntactic structures the utterances contained and the phonological alternations that some of the syntactic constituents undergo, e.g. contraction of function words. I also tried phonological explanation of general rules of assimilation and reduction, in addition to constant exposure to the sound of the words and connected forms; but nothing changed. Finally, I turned to pronunciation drills and listening discrimination and identification tasks were attempted. In the discrimination task, learners were asked to discriminate between two utterances and report whether the utterances were the same or different. In the identification task, learners were trained on the sound of the utterances and in a subsequent exercise were asked to identify them. Regardless of the type of technique and frequency and duration of exposure to any one technique, the identification of connected utterances was rarely realized.

I also made sure that culture and pragmatics information was provided as means to finding the necessary clues to help the learners recognize the components of the utterance. For example, listening for purpose, or pre-listening exercises that trained learners in strategic listening, such as listening with expectations and listening for the gist, were carried out. Although these exercises helped tremendously the learners in the process of parsing the overall meaning of the heard message, they did not help the learners to identify the components of an utterance.

Finally, I asked the learners to write immediately and as exactly as they could the content of the utterances they could hear and parse in prerecorded utterances. The reports showed that learners could not distinguish whether the utterance was made of a number of words or only one word. For example, they parsed the following utterances as constituents made of a single word or two. As seen, some utterances were parsed as one word as in (4b & c) or of two words as in (4a & c).

4) English input   CA speaking output
    a. Does it snow there?
        dʌzɪtˈsno ðɚ   ʔɪts nover

                                                ʔɪz noɡar

    b. Is it raining too?
        ɪzɪtˈreinɪŋ tu   bɨre:nɪngɨtu

                                                ʔɪzɡɨrɪrlitu

    c. How do you come to know that?
        hawʒʲəˈkʌm təno θæt   he:kinoni tumi

                                                ʔaɪtikʌmtuze

                                                hertikʌmtuzat

Intriguingly, when the learners were provided with the text versions of the oral discourse they heard, they were amazed that they could not parse the words. They insisted that what they heard did not correspond at all, to what they read.

A contrastive analysis between the phonology of English and that of CA was warranted particularly since a similar learning problem is found to exist among English speaking learners of Arabic. For example, Broselow (1979) found that American English speaking learners could not segment certain phrases of CA as illustrated in (5). They tend to

segment the connected utterance at [bintis], which is lexically nonexistent in CA, and then the string [mina]. In other words, they segment the phrase into two distinct phonological words. The phonological word is a prosodic element in the prosodic hierarchy (see the Prosodic Structure in 7 below).

In CA connected speech, CA follows strict limitations based on the well-formedness of syllables. Syllables can only be drawn from the following set: CV, CVV, CVC or phrase-final CVVC or CVCC. CA does not allow a cluster of three consonants; therefore, if through concatenating and connecting words, such a cluster is generated, an epenthetic vowel is inserted (Broselow, 1979, 1988) as in (5).

5)   Words in isolation   Words concatenated

a. bint simi:na   →   bintɨsmi:na

a fat girl

b. katábt gawá:b   →   katábtɨgawá:b

you wrote a letter

Based on an unpublished study conducted by the writer (Aquil, 1999), vowel insertion is also routinely observed in the learners' output when they were asked to repeat English connected speech. The data in (6) shows that learners do not recognize English words (e.g., 6 a, b, c, d), and parsed them differently (e.g., 6 e). They also inserted vowels in places that did not have any vowels in the corresponding English words (e.g., 6 f, g).

6)   English input   CA output
a. Oh, it's not so bad outside

ðis werro bai aʊtsaid

ɹes ɹu non kɨlin bara ʔaʊtsaid
ernotɨbadi ʔaʊtsaid

b. I don't like dust

ʔabiwain dʌst

ʔarlɨ̠ju laik ðis

This matter needs to be resolved

ðiz simail ɹu
ðiz mara niːd

ʔiθ narit ʔiðiz jiːr

c. He is always on time

noʔon hon ʔontaim
wazint wʌz taim

huzon ʔentaim

d. He has a lot of money

ʔizɨ̠laramʌni

ʔizra<u>nim</u>ʌni

a. Is she in this class?

romim <u>kɨ̠l</u>aas

ʔiʃi ʔiz <u>kɨ̠l</u>aas

ʔiʃ ʔiz <u>kɨ̠l</u>aas

ʔiʃi za <u>kɨ̠l</u>aas

b.   Is it raining too?

ʔiza bɨreːningitu

ðə bɨriŋ ton

bɨreːningitu

ʔizgɨrɪrlitu

neneːjɨngitu

Employing epenthesis to break up consonant clusters has been studied cross linguistically and researchers found that speakers of different languages insert vowels between consonant clusters of English words (Alber, 2000; Alber & Plag, 1999; Alsin & Pisoni, 1980; Asci, 1996; Broselow, 1999, 2001; Broselow, Chen, & Wang, 1998; Carlisle, 1991; Hansen, 2001, 2004; Kim & Jung, 1998).

Works by Broselow on Arabic (1979, 1988) and the study conducted by the writer (Aquil, 1999) suggest that English and Arabic employ contrasting phonological processes in connected speech. For example, English employs deletion, assimilation, and reduction and CA employs epenthesis. In English, through compression the pronunciation of the unstressed syllables changes, resulting in a surface output which is considerably different from the individual citation forms (Dalby, 1986; Selkirk, 1984, 1996; Zsiga, 2003; Zwicky, 1972). On the other hand, in CA, through resyllabification to obey syllable structure rules, vowels are sometimes inserted, (Harrell, 1957; McCarthy, 1979; Mitchell, 1956; Youssef & Mazurkewich, 1998). The change in the pronunciation resulting from such operations, that is, resyllabification, and vowel insertion, may not result in such a distinct and even opaque output as that found in connected speech of English. However, as mentioned above, English speakers learning Arabic have also reported difficulty in segmenting utterances that have undergone resyllabification (Broselow, 1988). It would

seem that this correspondence is not accidental; there may be something that both languages either share or do not share that causes the same problem for those who are learning the other language. To my knowledge, no studies from the perspective of Arabic speakers have looked in what may be the cause of learners' problems in listening.

Therefore, a contrastive analysis between the phonology of English and that of CA is warranted, particularly an analysis of the prosody of these two languages. It seems that these listening problems could be attributed to either similarities or differences between the prosodic structure of both languages.

2. Contrastive analysis between CA and English

2.1. Language transfer and contrastive analysis

It is well established that first language interference plays a major role in Second Language Acquisition (SLA). In 1957, Lado introduced the Contrastive Analysis Hypothesis (CAH), which called for a contrastive analysis between the grammars of the native language (L1) and that of the target language (L2) in order to predict learners' errors. It postulates that areas that are similar between the two languages will not pose any difficulty or generate errors, whereas areas that are different will be difficult and hence produce errors (Lado, 1957). However, empirical studies of transfer were few in the 1960s and the predictive validity of many contrastive analyses was found to be questionable (Odlin, 1989). In the 1970s and 1980s empirical research raised the question about the source of L2 errors (Duskova, 1983; Kellerman, 1977; Kellerman & Sharwood-Smith, 1986; Selinker, 1972; Zobl, 1980) and highlighted the importance of development errors in both L1 and L2 acquisition. This type of errors is not only common among L2 speakers of different language backgrounds but also among children learning their native language (Dulay & Burt, 1973, 1974).

Nevertheless, the belief among linguists that language interference must have some role in the acquisition process persisted. In the 1980s empirical studies were conducted, learners' language was investigated in detail, and performance data was

systematically gathered and analyzed. Language transfer was revisited providing evidence that elements from one language could be incorporated into another (Kellerman & Sharwood-Smith, 1986) and that the transfer of these elements was at all levels of the language, especially at the level of the sound system (Ellis, 1994, 1997).

Agreeing with the original CAH, Odlin (1989) adds that there are two types of transfer, negative and positive transfer and that both types of transfer involve similarities present between L1 and L2. Contrary to how similarities initially were viewed (cf. Lado, 1957), Oldin considers similarities as facilitators at times and inhibitors at other times. Positive transfer occurs when similarities between L1 and L2 help in the acquisition process. As an example, Odlin refers to similarity and cognates in vocabulary between L1 and L2, which can reduce the time needed to develop reading comprehension. Similarity in phonemic systems between L1 and L2 has a positive effect on a learner's discrimination between vowels. The other type of transfer is negative transfer, which involves divergences from norms in the target language. Negative transfer as manifested in learner' errors, occurs either due to similarities or differences between the native and target language and can take the form of: (a) substitutions, which involve the use of a native language structure in the target language, (b) calques, errors which reflect native language structure as evident in idiomatic and word order errors, and (c) alternations of structures, and hypercorrections, cases of overreactions to a particular influence from L1.

Odlin (1989) states that the influence from L1 results from a learner's conscious or unconscious judgment that linguistic elements in the native and target language are similar. Native language phonology can influence the interpretation of the target language. For example, misperceptions of the sounds of the target language are likely to be categorized in terms of the native language phonology (See Goto, 1971; MacKain et al., 1981 ; Miyawaki et al., 1975; Mochizuki, 1981) on Japanese learners' difficulty in perceiving English /l/ and /r/.

2.2. Similarities between English and CA

English and CA are both classified as stress-timed languages (Hayes, 1995; McCarthy, 1979; Watson, 2002). Rhythmically, languages can be classified as stress-timed or syllable-timed languages (Pike, 1945). Stress-timed means that stress occurs at approximately equal intervals, which could contain different numbers of syllables. In order for stress to occur at equal intervals, languages take different measures to fulfill this requirement. In some languages, the time allotted to syllables is reduced or shortened, whereas in other languages, time that is allotted to syllables is spread. In the case of English, the former applies, but with CA the latter apparently may take place. In English, words are contracted and consonants are blended, however, not at the expense of stress. A stressed syllable is respected and hence never reduced or deleted. English shows a great deal of phonetic reduction, especially of function words. A phonological rule prevails in English, namely "monosyllabic destressing", which applies to satisfy "rhythmic restructuring" (Selkirk, 1984). In fast speech, function words that are monosyllabic and carry weak stress undergo some phonological deletion processes of vowels and sometimes of certain consonants. For example, in (2a) [wʌʒʲəwanə it]
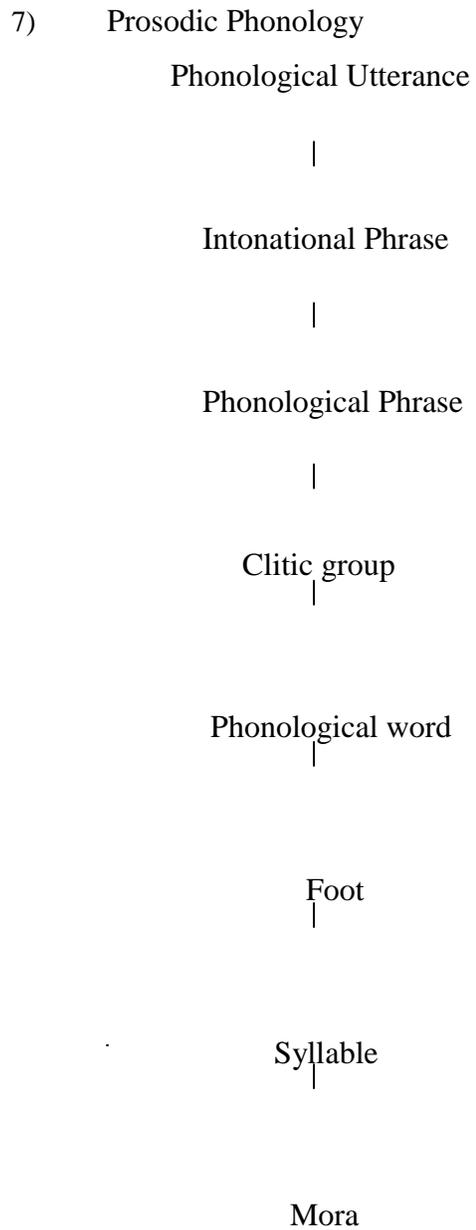
the monosyllabic function words *do*, *you* undergo vowel reduction and then deletion. Then, consonants blend and assimilate with the final consonant of the question word *what* resulting in a palatalized allophone [ʒʲə] (Selkirk, 1984: 344). On the other hand, in Arabic in general and CA in particular, when word concatenation takes place, it usually occurs in a way that would not result in an illicit syllable. That is, postlexical resyllabification always preserves the original segmental content by inducing insertion instead of deletion (cf. data in 5 above). Hence, both languages respect stress in connected spoken speech, however, each language utilizes a different tool, one uses contraction and blending, e.g. English and the other uses epenthesis, e.g. CA.

Psycholinguistic studies conducted on L1 segmentation units found that languages differ in the segmentation units being exploited in segmenting connected speech.

Segmentation is an abstract process. It is likely to help in locating and recognizing words in speech stream. Through this abstract process the listener segments the speech based on the specifics of his language prosody. Research in psycholinguistics found that languages differ in the way prosodic properties of each language are mirrored in the segmentation procedure adopted. (Cutler, 1997; Otake & Cutler, 1996; Otake, Hatano, & Yoneyama, 1996). For example, English uses the foot as a unit, where French uses the syllable (Cutler et al., 1986), Japanese the mora (Cutler et al., 1997; Cutler & Otake, 1994; Otake, Cutler, & Mehler, 1993; Otake et al., 1996) and Arabic the stressed syllable (Aquil, 2011, 2012).

By employing a syllable monitoring technique in a cross-modal design, (visual and auditory) Aquil (2012) finds evidence that a stressed (CVC) could be one of the preferred prosodic units used in the segmentation of CA connected speech. Subjects' recognition and miss rate of a stressed (CVC) syllable is quicker and more accurate in comparison with a (CV) syllable. Also in another psycholinguistic study, Aquil (2011) using word spotting technique finds evidence that stressed syllables are recognized faster and more accurately than any other syllable.

The syllable and the foot (i.e. stress) are constituents in the prosody of languages. According to Nespor and Vogel (1986), the prosody of a language is structured in a hierarchical way, with constituents both below the word level (e.g., the mora, the syllable, and the foot) and above the word level (e.g., the phonological phrase, intonational phrase and phonological utterance). The structure is illustrated in (7).

7)   Prosodic Phonology

Phonological Utterance

|

Intonational Phrase

|

Phonological Phrase

|

Clitic group
|

Phonological word
|

Foot
|

Syllable
|

Mora

In this paper, I focus on the syllable and the foot (i.e., stress) and the process of syllabification in CA and English. As noted in the previous section on segmentation, these two prosodic units have been the points of investigation in English and French and Cairene Arabic. As for postlexical phonological changes such as assimilation, consonant blending, vowel contraction and deletion, they will not be discussed in this paper due to space consideration and will be analyzed and discussed in subsequent papers.

2.3. Differences between English and CA

The syllable structure in CA and English

Syllable structures in both of CA and English are different. For example, CA syllable structure contains an obligatory onset, a nucleus, and a coda, which is optional. English syllable structure, on the other hand, includes not only optional onsets and codas but also consonants clusters in both of them, as illustrated in the following tables.

Table 1: Syllable Structure in English and CA

Eng   CA

σ   σ

ons   nuc   cod   ons   nuc   cod

(c..)   v   (..c)   c   v   (c)

English allows syllables to start without an onset. Consider the following table.

Table 2: Syllable types in Eng and CA

Eng		CA		
v	a	----		
cv	be	cv	bi	with
ccv	pry	----		
cccv	stray	----		
vc	oat	----		
cvc	men	cvc	bes	enough
ccvc	class	----		
cccvc	stroke	----		
vcc	apt	----		
cvcc	land	cvc(c)	bint	girl
----		cvv	fii	in
ccvcc	shrilled	-----		
cccvcc	string	-----		
vccc	angst	-----		
cvcc	range	-----		
ccvccc	sphinx	-----		
cccvccc	scrounged	-----		

| vcccc | angsts | ----- |
| cvccccc | texts | ----- |

cvv(c) ħu:t   whale

The data and the structures demonstrate that English has a wider range of syllables than CA, and can start a syllable in a vowel , whereas CA has an obligatory onset and allows consonant clusters (CC), however; limits their distribution to word final only.

Stress in CA and English

Description of stress placement in CA was first formulated by McCarthy (1979) and later reported in Halle and Vergnaud (1987), and Hayes (1995). It is as follows (Halle and Vergnaud, 1987, p. 61) cited in Youssef & Mazurkewich (1998):

8) Stress is on the last syllable if it is superheavy:

   a. darábt   I/ you beat

   b. haggáat   pilgrimages

9) Otherwise, stress is on the penult, if it is heavy:

   a. béetak   your (m.sg) house

   b. katábti   you (fem, sg) wrote

10) Otherwise, stress is on the antepenult or the penult, whichever is separated by an even number of syllables from the immediately preceding heavy syllable if there is one or the beginning of the word if there are none, where zero separation is counted as even:

   a. katabítu   she wrote it

   b. muxtálifa   different (adj. s.f.)

   c. šágara   tree

On the other hand, English stress rules are more complex than those in CA, primarily because stress placement in English interacts with morphology and syntax (see data in (11). In example 11)a) the weight of the final syllables attracts stress, e.g., tense vowel in i) and consonant cluster in ii). If the final syllable is not heavy, stress is on the penult. In nouns as well, weight of a syllable attracts stress as can be seen in example (11b). As for adjectives, stress depends entirely on whether or not the adjective contains a suffix. Unsuffixed adjectives follow the same pattern as that of verbs (Data adapted from Chomsky & Halle, 1968; Halle & Vergnaud, 1987; Youssef & Mazurkewich, 1998).

11)

    a. In verbs stress is assigned to:

        i. Final syllables with tense vowels

           volunteer

           achieve

        ii Final cluster of two consonants

           collapse

           direct

        iii. Otherwise the penultimate syllable

           astonish

           edit

    b. In nouns stress is assigned to

        i. Penultimate if it is heavy

    .   veranda

consensus

agenda

appendix

.   ii. Antepenultimate if the penult is light

america

cinema

javelin

venison

This is a very simplified account as it is well documented that there are counter examples to all of the above stress placements. Although analysis of these counter examples is of absolute importance, it is beyond the scope of the present study to carry out such an analysis at this point.

Syllabification in CA vs. English

According to Nespor & Vogel (1986), the domain of syllabification in English is the phonological word. Syllabification does not occur across the members of a compound, nor does it apply across words in a phrase. Nespor and Vogel (1986) gave evidence by demonstrating that the [k] in (12) could not be the onset of the second syllable and was in fact the coda of the first.

12)

    a.   pecan   [pe] σ[can] σ   [pi]σ [khæn] σ

    b.   pique Anne   i.   [pique] σ[anne]   [pik] σ[æn] σ

                  ii * [pi[k]] σ[[k]anne]   *[pi] σ[khæn] σ

c. pack ice (N)   i   [pack] σ[[ice] σ   [pæk] σ [ais] σ

                  ii   *[pa] σ[[k]ice] σ   * [pæ] σ [khais] σ

d. pack ice (VP)   i.   [pack] σ [[ice] σ   [pæk] σ [ais] σ

                  ii. *[pa] σ [[k]ice] σ-* [pæ] σ [khais] σ

                                    (p. 63, 67)

More evidence that syllabification does not occur across words in a phrase in English, Nespor and Vogel cite an example where the pronunciation of a segment differs according to its position in the syllable. The segment [l] is a clear [l] when it is initial but dark [ɫ] when it is final.

13)

      [l]ap   clear

      pa[ɫ]   dark

They add that if syllabification occurred across words in English, one would expect a dark [ɫ] to become clear [l] in *call Andy*. On the contrary, resyllabification does not take place because the [l] of *call* stays dark [ɫ] and it is very different from a clear [l] as in *Yolanda* (Nespor & Vogel, 1986, p. 65).

On the other hand, syllabification and resyllabification in CA as seen below occurs across words, and in fact is obligatory across words if a sequence of more than two consonants result due to morpheme or word concatenation as the data in (14) demonstrates.

14)   ka táb t lu   →   [katábtɨlu]

                                          'I wrote to him'

        ha+ tikuun   →   [hatkúun]

                                          'you (m.sing) will be'

        kálb +na   →   [kalbɨ́na]

                                          'our dog'

        ʕamr batt   →   [ʕamrɨbatt]

                                          'ʕamr decided'

        ʕamr batt fi xusaarit ?imbaariħ →   [ʕamrɨbattɨfɨxsartimbaariħ]

                                          'ʕamr decided in yesterday's loss'

## 3. Solutions to the learning problem

The solutions proposed in this paper arise from the contrastive analysis conducted on the syllable structure and stress in English and CA. The analysis demonstrates that syllable structures are different. One of the most important aspects related to CA syllable structure is that syllables cannot start with a vowel. Therefore, a theoretically based solution related to this aspect is provided below.

The analysis also illustrated that stress can be another aspect that could pose problems for CA speaking learners of English. Therefore, an applied solution related to stress is presented below.

Theoretical solution

Syllable structure and developmental stages

Based on the comparison between syllable structures in English and CA (see Table 1 & 2) and based on universal implications (Eckman, 2004; Greenberg, 1963, 1987) and developmental stages in language acquisition (Pienemann, 1984, 1989, 1998), I propose that the vital step in the acquisition of English syllable structure for a CA learner is to gain knowledge that a syllable can be onsetless as in [V/VC] as demonstrated in Figure 1.
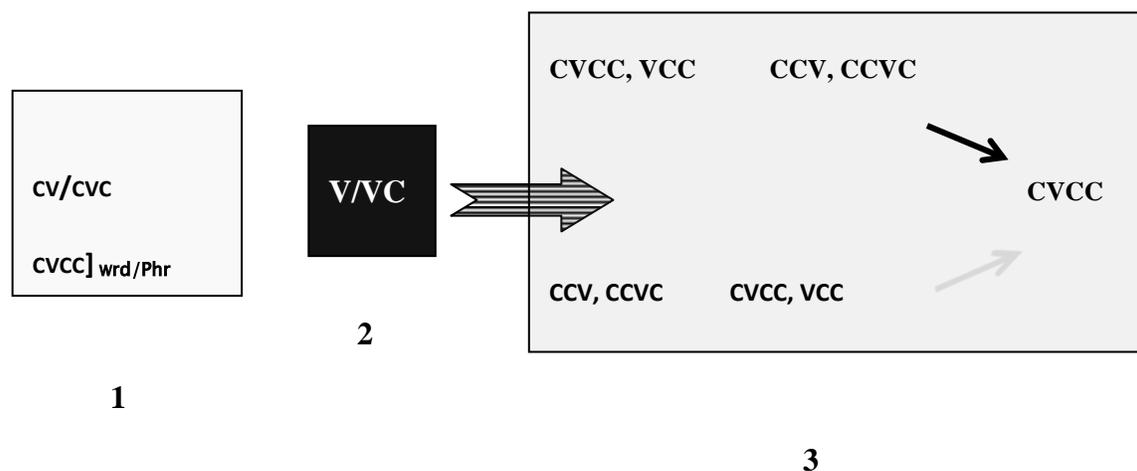
Figure 1: Stages of syllable structure Learnability by CA learning English L2

Adapted from Morales (1999: 9)

Although Morales' (1999) stages of acquisition are for First Language Acquisition, I would like to adopt these stages to the CA L2 learner since the similarity and correspondence between FLA and SLA has been established in the literature.

Based on universal implications (Eckman, 2004; Greenberg, 1963, 1987) and developmental stages in language acquisition (Pienemann, 1984, 1989, 1998) by teaching a CA learner first the structures of VCC and CCV the acquisition of V, VC, CCVC, and CCV English syllable structures could be acquired, as they are implied in the structures of VCC and CCV.

A CA learner of English comes in with a grammar that has CV/CVC and CVCC/CVVC word and phrase final (see Table 2). By teaching him that English has syllables and words, which start with a vowel (see Table 2) as illustrated with box 2 in figure 1, the remaining structures will be learned by implication. In other words, the focus should be on teaching the distinct English syllable structure as in box 2. This is by giving examples of words like *'apple',* and *'oat'* in citation forms and in connected speech. Needless to mention, is that learners should not be inundated with linguistic rules. The point here is to make them aware of the English distinct syllable structures and the syllabification process that each language employs (see section 2.3.3 above).

Applied solution

Segmentation as a technique

The solution proposed here adopts Field (2003) prosodic segmentation to replace lexical segmentation, where learners process spoken language to recognize single words presented in a speech stream or in their canonical citation forms (Grosjean & Gee, 1987).

Field asserts that L2 listeners with weak listening skills tend to adopt a strategy that is based on lexical or vocabulary extraction, by which the learner searches in the signal for matches between sequences of sounds and items of known vocabulary. In anxiety and concern to locate matches, word boundaries are often breached and ignored. Learners need to be made aware of the tentative nature of the matches they find in the signal and instructed not to immediately construct a mental scenario of the text based on

the provisional match they might spot in the signal. For example in the following situation, student 1 could have erroneously constructed a complete mental script based on the inaccurate extraction.

15)

Speaker: went to assist a passenger   Student 1 extracts: sister

Speaker: the standard the hotel achieves Student 2 extracts: stand at the hotel
(Adapted from Field, 2003:328)

To replace lexical segmentation, Field (2003) suggests adopting native speakers' prosodic or metrical segmentation strategy. Field's method concentrates on the more salient syllables which usually start English content words and often initiate a lexical search. For example in the case of English as a second or foreign language, explain that spoken language is made of prosodic units such as strong and weak syllables, the distribution of which is language specific. Train the learners to notice the distribution and perceptual differences between these two types of syllables. Ask them to identify and distinguish between these units, especially since spoken language is not made of only content words represented in strong syllables but also function words that as important as content words. These functions words are usually reduced and are not as salient as content words. Make them aware that confusion on function words may lead to slips of the ear. Listeners are likely to deal with a preposition following a content word as the second syllable of the word as Grosjean has found in his gated experiment (Grosjean, 1980: 139).

16)   Bun in   ⟶   bunny

      boar in   ⟶   boring

      plum on   ⟶   plumber

One way to develop saliency as a cue for content words is to use the technique described by Field (2003). This technique recommends playing recordings at a low volume. Then ask the learners to transcribe first the more salient syllables. Stressed syllables are 'islands of reliability', since they are usually louder and longer than the unstressed ones. For a second turn, let them pay attention to the distribution of stressed vs. unstressed syllables and ask them to focus their listening on weak syllables surrounding strong syllables or content words. Have them identify these function words, but make use of all learners' listening or guessing errors and explain to them that by perceiving a preposition as a second syllable of a word can yield a different interpretation and hearing (e.g. example in 16).

Focusing the learners' attention on stressed syllables, and the distribution of stressed versus unstressed ones, may help them not only locate word boundaries, but also become aware that recognizing the pattern is as important as spotting the lexical word.

Acquaint your learners with the linguistic fact that alternations in boundaries between words in a spoken language could yield different understandings of a text. Boundary alternations like, boundary shifts, deletions and substitutions as in the following data (Grosjean & Gee, 1987: 140).

17)   an ice bucket   ⟶   a nice bucket

ten year party   ⟶   tenure party

descriptive linguistics   ⟶   the script of linguistics

Make learners aware of the need for caution in placing, shifting, or deleting boundaries, by making them listen to ambiguous sequences where more than one meaning could arise, and then disambiguate the sequence by adding more words which in turn could result in changing the meaning all together. This can be done by either dictation, or by having the learners listen to a recording where ambiguous sequences are

followed by a pause. During this pause, the learners repeat the sequence either orally or in writing. Then the disambiguating sequence follows to demonstrate how inserting, shifting or deleting word boundaries in different locations could result in different interpretations of an auditory text. Consider the following example adapted from Field (2003:328).

18)

   a. Teacher: an ice cream ….. [Ss write] …. Teacher then continues: a nice cream dress

   b.   Teacher: the boxes of ……[Ss write]….. Teacher then continues: the box

     have been opened.

Provide anticipatory or remedial practice on determining where word boundaries may fall especially in postlexical phonological alternations such as in contraction, vowel reduction, resyllabification and cliticization of function words in English.

Finally, demonstrate to learners that they also could fall victim to erroneous segmentation in their own native language due to boundary alternations. For instances the examples from CA below illustrate that a shift in word boundary can in fact result in changing the meaning, sometimes drastically.

19)   Ɂabúu našálha   Ɂabúuna šálha

       his father stole it (f)   our father carried it (f)

       Ɂabúu naɁálha   Ɂabúuna Ɂálha

       his father moved it (f)   our father said it (f)

       kal bi šaalu   kalb ɪ šaalu

he ate with his shawl	a dog killed him (lit., carried)
gi-d-dok tˤoor	giddak tˤoor
the doctor came	your grandfather is a beast (lit., bull)

Conclusion

As seen CA and English, belong to the same language classification, namely, stress-timed language. Both languages use stress as a segmentation unit. Hence, the similarities that might inhibit CA speaking learners' parsing English connected speech. L2 language pedagogy has always concentrated on teaching learners the differences rather than the similarities. However, as seen in this paper a negative language transfer could take place based on the similarities. Based on the data and information provided in this paper, English L2 pedagogy in the Arab world may need to take into account and draw learners' attention to both the differences and similarities between Arabic and English especially in the syllable structure and stress systems. A theoretical and applied solutions are offered in this paper. The theoretical solution advocates making learners aware of the English distinct syllable structure V/VC, whereas the applied provides techniques to train learners how to segment English connected speech based on the distribution of stressed and unstressed syllables. In conclusion, further empirical studies investigating pedagogical interventions based on the proposed theoretical and applied solutions are warranted and recommended.

About the author:

Rajaa Aquil, Assistant Professor of Arabic at Georgia Tech, Atlanta, Georgia, USA. Her research and work is on psycholinguistics, specifically spoken languages. She published studies in the Perspectives or Arabic linguistics and in the Journal of Psycholinguistic research. In foreign language pedagogy, she focuses and works on L2 listening skill.

References

Alber, B. (2000). Regional variation at edges: Glottal stop epenthesis and dissimilation in Standard and Southern varieties of German, from http://roa.rutgers.edu/view.php3?roa=417

Alber, B., & Plag, I. (1999). Epenthesis, deletion and the emergence of the optimal syllable in creole, from http://roa.rutgers.edu/view.php3?roa=335

Alsin, R. N., & Pisoni, D. B. (1980). Some developmental processes in speech perception. In G. Yeni-Komshian, J. Kavanagh & D. Ferguson (Eds.), *Child phonology: perception and production: Vol 2, perception* (pp. 67-96). New York: Academic Press.

Aoyamaa, K., Flege, J. E., Guiona, S. G., Yamada, R. A., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics, 32*, 233-250.

Aquil, R. (1999). *Focus on Form as an Instructional Procedure in Solving Segmentation Problems for Listeners*. Qualifying paper. Linguistics. Georgetown University.

Aquil, R. (2011). Empirical Evidence for Stress as a Perceptual Unit in Spoken Cairene Arabic. *Perspectives of Arabic Linguistics, 23rd & 24th* 3-20.

Aquil, R. (2012). The role of the syllable in the segmentation of Cairene spoken Arabic. *Journal of Psycholinguistic Research, 41*(2), 141-158.

Asci, A. (1996). Consonant clusters and epenthetic vowels in Turkish. *Travaux de l'Insitut de Phonetique de Strasbourg, 26*, 1-31.

Broselow, E. (1979). Cairene Arabic syllable structure. *Linguistic Analysis, 5*(4), 345-382.

Broselow, E. (1988). Prosodic phonology and the acquisition of a second language. In S. Flynn & W. O'Neil (Eds.), *Linguistic theory in second language acquisition* (pp. 295- 308): Kluwer Academic Press.

Broselow, E. (1999). Stress, epenthesis, and segment transformation in Selayarese loans.

Broselow, E. (2001). Stress-epenthesis interactions, from http://roa.rutgers.edu/view.php3?roa=446

Broselow, E., Chen, S. I., & Wang, C. (1998). The emergence of the unmarked in second language phonology. *Studies in Second Language Acquisition, 20*, 261-280.

Carlisle, R. S. (1991). The influence of environment on vowel epenthesis in Spanish/English Interphonology. *Applied Linguistics, 12*, 76-95.

Chen, Q. (2000). *Analysis of Mandarin tonal errors in connected speech by English-speaking American adult learners: A study at and above the word level.* Ph.D. Dissertation, Brigham Young University.

Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper & Row, Publishers.

Cutler, A. (1997). The comparative perspective on spoken language processing. *Speech Communication, 21*, 3-15.

Cutler, A. (1999). Prosodic structure and word recognition. In A. D. Friederici (Ed.), *Language comprehension: A biological perspective* (Second Edition ed.). New York: Springer.

Cutler, A., & Butterfield, S. (1992). Rhythmic cues of speech segmentation: Evidence from juncture misperception. *Language and Cognitive Processes, 31*, 218 - 236.

Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: Literature review. *Language and Speech, 40*, 141-201.

Cutler, A., & Norris, D. (1988). The role of strong syllable in segmentation for lexical accesss. *Journal of Experimental Child Psychology: Human Perception and Performance, 14*, 113-121.

Cutler, A., & Otake, T. (1994). Mora or phoneme? Further evidence for language-specific listening. *Journal of Experimental Psychology: Human perception and Performance, 14*(113-121).

Dalby, J. (1986). *Phonetic structure of fast speech in American English.* Ph.d dissertaion, University of Indiana: Indiana University Linguistics Club, Bloomington, Indiana.

Dulay, H., & Burt, M. (1973). Should we teach children syntax? *Language Learning, 23*, 245-258.

Dulay, H., & Burt, M. (1974). You can't learn without goofing. In J. Richards (Ed.), *Error analysis*. London: Longman.

Duskova, L. (1983). On sources of erros in foreign language learning. In B. Robinett & J. Schachter (Eds.), *Second language learning: Contrastive analysis, error analyis, and related aspects* (pp. 215-233). Ann Arbor: University of Michigan Press.

Eckman, F. (2004). From phonemic differences to constraint rankings. *SSLA, 26*, 513-549.

Ellis, R. (1994). *The Study of second language acquisition*. Oxford: Oxford University Press.

Ellis, R. (1997). *Second language acquisition*. Oxford: Oxford University Press.

Flege, J. E., Bohn, O. S., & Jang, S. (1997). The effect of experience on nonnative subjects' production and perception of English vowels. *Journal of Phonetics, 25*(169-186).

Flege, J. E., & MacKay, I. R. A. (2004). Perceiving vowels in a second language. *SSLA, 26*, 1-34.

Flege, J. E., MacKay, I. R. A., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *Journal of Acoustical Society of America, 106*(5), 2973-2987.

Frauenfelder, U., & Floccia, C. (1999). The recognition of spoken words. In A. D. Friederici (Ed.), *Language comprehension: A biological perspective* (Second Edition ed.). New York: Springer.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sound "L and "R". *Neuropsychologia, 9*, 317-323.

Greenberg, J. H. (1963). *Universals of language*. Cambridge, MA: MIT Press.

Greenberg, J. H. (1987). Some generalization concerning initial and final consonant clusters. In J. Greenberg, C. Ferguson & E. Moravckik (Eds.), *Universals of human language* (Vol. 2, pp. 243-279). Stanford, CA: Standford University Press.

Grosjean, F. (1985). The recognition of words after their acoustic offset: Evidence and implications. *Perception and Psychophysics, 38*(299-310).

Grosjean, F., & Gee, J. P. (1987). Prosodic structure and spoken word recognition. In U. Frauenfelder & L. K. Tyler (Eds.), *Spoken word recognition*: A Bradford Book. The MIT Press.

Guiona, S. G., Flege, J. E., Yamada, R. A., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of Acoustical Society of America, 107*(5), 2711- 2724.

Halle, M., & Vergnaud, J. R. (1987). *An Essay on Stress*. Cambridge, Massachusetts: The MIT Press.

Hansen, J. G. (2001). Linguistic constraints on the acquisition of English syllable codas by native speakers of Mandarin Chinese. *Applied Linguistics, 22*, 338-365.

Hansen, J. G. (2004). Developmental sequences in the acquisition of English L2 syllables codas. *Studies in Second Language Acquisition, 26*, 85-1245.

Harrell, R. S. (1957). *The phonology of colloquial Egyptian Arabic*. New York: American Council of Learned Societies.

Hayashi, T. (1991). Interactive processing of words in connected speech in L2 and L2. *IRAL, 29*(2), 151-160.

Hayes, B. (1995). *Metrical stress theory: Principles and case studies*. Chicago: University of Chicago Press.

Kellerman, E. (1977). Towards a characterization of the strategy of transfer in second language learning. *Interlanguage Studies Bulletin, 2*(58-145).

Kellerman, E., & Sharwood-Smith, M. (1986). *Cross-Linguistic Influence in Second Language Acquisition.* Elmsford, NY: Pergamon.

Kim, D.-Y., & Jung, Y. (1998). Adult ESL Korean speakers' interlanguage acquisition of consonant clusters. *The SECOL Review, 22*(2), 150-169.

Ladefoged, P. (2001). *A course in phonetics: Fourth edition*. USA: Heinle & Heinle. Thomson Learning.

Lado, R. (1957). *Linguistics across cultures*. Ann Arbor, MI: The University of Michigan Press.

Lenhardtova, L. (1993). An experimental study of phonological interference in the English of Slovaks. *IRAL, 31*(4), 291-308.

MacKain, K., Best, C. T., & Strange, W. (1981). Categorical perception of English /r/ and /l/: a first report. *Journal of the Acoustical Society of America, 89*(2), 874-886.

McCarthy, J. (1979). On stress and syllabification. *Linguistic Inquiry, 10*, 443-466.

McQueen, J. M., Norris, D. G., & Cutler, A. (1994). Competition in spoken word recognition: Spotting words in other words. *Journal of Experimental Psychology: Learning, Memory and Cognition, 20*, 621-638.

Mehler, J., Dommergues, J., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior, 20*, 298-305.

Mitchell, T. F. (1956). *An introduction to Egyptian coloquial Arabic* New York,: Oxford University Press.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A. M., Jenkins, J., & Fukimura, O. (1975). An effect on linguistic experience: The discrimination of [r] and [l] by native speakers of Japanese and English. *Perception & Psychophysics, 18*(5), 331-340.

Mochizuki, M. (1981). The identification of /r/ and /l/ in natural and synthesized speech. *Journal of Phonetics, 9*, 283-303.

Morales, A. F. (1999). *Stages in the acquisition of syllable typology*. Georgetown University.

Nespor, M., & Vogel, I. (1986). *Prosodic phonology*. Dordrech-Holland/Riverton- USA: Foris Publications.

Odlin, T. (1989). *Language transfer: Cross-Linguistic Influence in Language learning*. Cambridge: Cambridge University Press.

Otake, T., & Cutler, A. (1996). *Phonological structure and language processing: Cross-linguistic studies*. New York: Mouton de Gruyter.

Otake, T., Cutler, A., & Mehler, J. (1993). Mora or syllable? Speech segmentation in Japanese. *Journal of Memory and Language, 32*, 358-378.

Otake, T., Hatano, G., & Yoneyama, K. (1996). Speech segmentation by Japanese listeners. In T. Otake & A. Cutler (Eds.), *Phonological structure and language processing: Cross-linguistic studies*. New York: Mouton de Gruyter.

Padilla, M. (2003). *English phoneme and word recognition by nonnative English speakers as a function of spectral resolution and English experience.* Ph.D. Dissertation, University of Southern California.

Pienemann, M. (1984). Psychological constraints on the teachability of languages. *Studies in Second Language Acquisition, 6*, 186-214.

Pienemann, M. (1989). Is language teachable? Psycholinguistic experiments and hypotheses. *Applied Linguistics, 10*, 52-79.

Pienemann, M. (1998). *Language processing and second language acquisition: Processability theory*. Amsterdam: Benjamins.

Pike, K. L. (1945). *The intonation of American English*. Ann Arbor: The University of Michigan Press.

Selinker, L. (1972). Interlanguage. *IRAL, 10*(209-231).

Selkirk, E. (1984). *Phonology and syntax: The relation between sound and structure*. Cambridge, Massachusetts: The MIT Press.

Selkirk, E. (1996). The prosodic structure of function words. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstraping from speech to grammar in early acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Sheldon, A., & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: evidence that speech production can precede speech perception. *Applied Psycholinguistics, 3*(243-261).

Strange, W., & Dittmann, S. (1984). Effects of discrimination training on the perception of /r-l/ by Japanese adults learning English. *Perception & Psychophysics, 36*, 131-145.

Tyler, L. K., & Frauenfelder, U. (1987). The process of spoken word recognition: An introduction. In U. H. Frauenfelder & L. K. Tyler (Eds.), *Spoken word recognition*. Cambridge, Massachusetts: A Bradford Book: The MIT Press.

Watson, J. C. E. (2002). *The phonology and morphology of Arabic*. Oxford: Oxford University Press.

Youssef, A., & Mazurkewich, I. (1998). The acquisition of English metrical parameters and syllable structure by adult native speakers of Egyptian Arabic (Cairene Dialect). In S. Flynn, G. Mortahardjono & W. O'Neil (Eds.), *The generative study of second language acquisition*. New Jersey: Mahwah.

Zobl, H. (1980). Developmental and transfer error: Their common bases and (possibly) differential effects on subsequent learning. *TESOL Quarterly, 14*, 469-479.

Zsiga, E. C. (2003). Articulatory timing in a second language: Evidence from Russian and English. *Studies in Second Language Acquisition, 25*, 399-432.

Zwicky, A. (1972). On casual speech *Papers from the eighth regional meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society.