An investigation into the Differences between English and Arabic Consonant and Vowel Sounds: 
A Contrastive Study with Pedagogical Implications

Dr. Walid M. Amer 
Associate Professor of Linguistics 
The Islamic University of Gaza: Palestine
Dr.walidm@yahoo.com

1. Introduction

How do human speech sounds differ from one language to another? Are there any two languages that use the same group of sounds? How many different sounds are there in the world's languages? What are the roles of sound production in each language? All of these questions will probably face anyone who is exposed to a second, or a foreign language. Man is capable of producing an infinite number of sounds. The science that covers all these points is called phonetics. It studies the human speech sounds in natural languages.

This study discusses the differences between Arabic and English sounds. It observes the differences between manner and place of articulation of consonant and vowel sounds in English and Arabic. It also touches upon pedagogical implications for teaching English and Arabic phonology.

1. Consonants

1.1 English consonants

“A consonant is formed when the air stream is restricted or stopped at same point between the vocal cords and the lips” (Todd: 1987:14). The outward flow of breath is obstructed in various ways by the organs of speech in the production of English consonant sounds. Sometimes the flow of air is stopped completely. Sometimes the stoppage occurs only partially, so that friction occurs. Sometimes the airflow is forced over the sides of the tongue or made to pass through the nose. The most appropriate way of describing a consonant sound is in terms of place and manner of articulation.

1.1.1 The place of articulation or point of articulation

When the air stream passes through the vocal tract, it is obstructed in some way. Therefore, the most important articulators that may cause obstruction are lips, teeth, vocal cords, and the hard palate (Ladefoged, 2001). When the air passes through the larynx, it comes up and out through the mouth or the nose. Most consonant sounds are produced by using the tongue and other parts of the mouth which determine the shape of the oral cavity, through which the air is passing. The terms used to describe many sounds are those which denote the place of articulation of the sound; that is, the location inside the mouth at which point the constriction takes place. The most common eight places of articulation will be presented below:

- Bilabial: 
These sounds are formed by both the upper and lower lips when they come together. 
They are the initial sounds of “pie, buy, my” represented by the symbols /p/ , /b/ and /m/. The /w/ sound that is found in the beginning of “way” is also bilabial.
- Labiodental:
These sounds are formed by the lower lip and the upper teeth; they are found in words like “friend, vie” when the lower lip rises until it nearly touches the upper teeth. These sounds are represented by the symbols /f/ and /v/.

- Dental:
Dental sounds are formed with the tongue tip behind the upper front teeth. These sounds are available in words like “there, thin” and they are represented by the symbols /ð/ and /θ/.

- Alveolar:
These sounds are formed as a result of having the front part of the tongue touching or pointing up to the alveolar ridge. They are the initial sounds in “top, dip, sit, zoo, nut”. They are represented by the symbols /t/ /d/ /s/ /z/ /n/. Other alveolar sounds are /l/ sound in the beginning of words like “lap” and /r/ sound at the beginning of “rip, right”.

- Alveopalatal:
To produce these sounds the tongue touches the roof of the mouth behind the alveolar ridge near the hard palate. Examples involving such sounds are: “ship, child, measure, jeep”. These are represented by the symbols /ʃ/ /tʃ/ /ʒ/ and /dʒ/ respectively.

- Velar:
Velars are the sounds produced using the back of the tongue and the soft palate. They are the sounds /k/, /ɡ/ and /ŋ/. These occur at the end of “hack, hag, hang”.

- Palatal:
The sound produced with the front of the tongue and the hard palate. We have only one English palatal sound which occurs at the beginning of the word “you”. This sound is presented by the symbol /j/.

- Glottal:
One sound is produced when the glottis is open, but there is no manipulation of the air passing out through the mouth. The sound produced is presented by /h/, which is the first sound in “who” and “whose”.

- Retroflex:
Ladefoged (2001:7) mentioned the sound that is produced with the tongue blade touching the alveolar ridge. It occurs initially in words such as “rye, row”. If this sound is pronounced at the end of words by the same speakers, we may also have retroflex sound like “air”. This sound is represented by the symbol /r/.

1.1.2 The manner of articulation or degrees of occlusion

The manner of articulation refers to the way the airstream is configured during the production of consonant sounds. As we have just seen, sounds of English are distinguished by their place of articulation. However, we need to differentiate between some sounds which we have placed in the same category. For instance, the sounds /t/ and /s/ are both alveolar sounds. How do they differ? They differ in the way
they are pronounced; that is, their manner of articulation, i.e. the manner of the air stream. There are different manners of articulation in English.

- Stop/plosive:
The consonants produced this way are called so because the air stream in the vocal tract is completely stopped at some point. The closure can be made by the two lips, producing the bilabial plosive /p/ and /b/. It can be made by the tongue pressing against the alveolar ridge, producing the alveolar plosive /t/ and /d/; and it can be made by the back of the tongue pressing the soft palate producing the velar plosives /k/ and /g/.

- Fricative:
The sounds here are produced as a result of incomplete closure at some point in the mouth. This manner of articulation is used in producing a set of sounds which include the labiodental fricatives /f/ and /v/, the dental fricatives /θ/ and /ð/, the alveolar fricatives /s/ and /z/, the alveopalatal fricatives /ʃ/ and /ʒ/, and the glottal fricative /h/.

- Affricate:
Affricates are a combination of sounds which start with complete closure then it is followed by a slow release of air with friction. It includes the alveopalatal affricates /tʃ/ and /dʒ/.

- Nasal:
Unlike other consonants, Nasals have a complete closure of the mouth. The velum is lowered and air stream is allowed to flow out through the nose producing the bilabial nasal /m/, the alveolar nasal /n/ and the velar nasal /ŋ/.

- Lateral:
Laterals also involve partial closure in the mouth, that is, the air stream is blocked by the tip of the tongue but allowed to escape around the sides of the tongue. /l/ sound is alveolar lateral sound.

- Glides (semi-vowels):
These sounds are vowel-like because they are made without closure in the mouth. Gildings occur in the beginning of a word or syllable (Todd 1987). Semi-vowels are the palatal glide /j/ and the velar glide /w/. In addition, the production of /r/ leads to an assumption that it is a semi-vowel.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>LD</th>
<th>D</th>
<th>A</th>
<th>AP</th>
<th>V</th>
<th>P</th>
<th>G</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop</td>
<td>P, b</td>
<td>t, d</td>
<td>k, g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>f, v</td>
<td>θ, δ</td>
<td>s, z</td>
<td>f, ʒ</td>
<td>h</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td>tʃ, dʒ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td>η</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lateral</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glide</td>
<td>w</td>
<td>j</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1.2 Arabic consonants

Arabic has twenty-eight letters and three notations; nominative, accusative and genitive cases. This gives thirty one sounds. Consonants are found to differ in their
place and manner of articulation. The following points will cover the articulation system of the Arabic consonants in terms of the point of articulation and the state of the air stream, i.e. place and manner of articulation.

1.2.1 The place of articulation or point of articulation

The points that are responsible for producing sounds are called articulators and the sounds are related to them. Therefore, points of articulation are formed by the movement of an active organ towards another constant organ (Masloh, 1980).

- **Bilabial:**
Bilabials are produced when the lips come together as in the initials of “baṭah بَطَه, madrasah مَدْرَسَة, wisam وَسَام” and they are represented by the symbols /b/, /m/, and /w/.

- **Labiodental:**
In Arabic there is only one labiodental sound. This is formed with the upper teeth and the lower lip. The sound is available in the initial of “fīna`ا فِيَنَا” and this sound is represented by /f/.

- **Interdental:**
These sounds are formed by the tongue between the upper and the lower teeth. They are the initial sounds of “θawr θَوْر, ḍalek ظَلِك, ḍ̣̣ efir ظَفِير” and they are represented by the symbols /θ/, /ḍ/, and /ḍ̣̣/.

- **Alveolar:**
These sounds are formed with the front part of the tongue on the alveolar ridge. Examples of the alveolar sounds are available in the words “rashid رَاشِد, Laken لَاكِن, nour نُور” and they are represented by the symbols /r/, /l/ and /n/.

- **Dental:**
They are sounds produced by the tongue tip behind the upper front teeth. They are initial of “ṭaleb طَالِب, tamshi طَمْشِي, domiah دُمْيَة, ḍabet ضَابِط, saif سَيْف, ṣabar زَابِر, zahir زَاہِر” and they are represented by the symbols /t/, /ṭ/, /d/, /ḍ/, /s/, /ṣ/, and /z/.

- **Palatal:**
Palatals are sounds which are pronounced by the front of the tongue and hard palate. They are found in the words “shams شَمْس, yawm يَوْم, jamil جَمِيل” and they are represented by symbols /ʃ/, /j/, and /ʒ/.

- **Velar:**
Sounds produced with the back of the tongue against the velum. Velars are found in the beginning of “xadem خَادِم, g hali غَلَي, kamel كَامِل” and they are symbolized by /x/, /g/, and /k/.
- **Glottal:**
  They are the sounds produced when the glottis is open and there is no manipulation of the air passing out through the mouth. The sounds appear in the words “حاناٍك” and “إنسان”, and they are represented by the symbols /h/ and /ʔ/.

- **Pharyngeal:**
  Pharynx is involved in the production of pharyngeal sounds with the back of the tongue and the pharynx. They appear in the words “دداداءن” and they are represented by the symbols /ḥ/ and /ʕ/.

- **Uvular:**
  There is only one Arabic consonant uvular, which is produced by the back of the tongue with the uvula. This sound occurs in the beginning of the word “قافلة” and it is represented by the symbol /q/.

1.2.2 **Manner of articulation or the state of the air stream**

Sounds in Arabic like the sounds of other natural languages are produced by the organs of speech, which we call place of articulation that give characteristics for every sound (Omar, 1981). However, we may face some sounds which have the same point of articulation. Thus, we can differentiate between them by looking at the state of the air stream, that is, their manner of articulation. For more explanation, consider the manner of articulating the Arabic consonants below:

- **Stop:**
  Arabic stop consonants are the result of a complete closure at some point in the mouth. The pressure builds up behind the closure, and then the air is suddenly released. /b/ is a bilabial stop, /t/ , /ṭ/ , /d/ and /ḍ/ are dental stops, /q/ is a uvular stop , /k/ is a velar stop.

- **Affricate:**
  If you combine a brief stopping of the air stream with an obstructed release, which causes some friction, the resulting sound will be the palatal affricate /ʒ/.

- **Fricative**
  Producing these sounds involves blocking the air stream and having the air pushed through the narrow opening. In Arabic we find the labiodental fricative /f/, the interdental fricatives /θ/ , /ð/ , the dental fricative /s/ , /ṣ/ , /z/ ;the palatal fricative /ʃ/, the velar fricatives /x/ , /ɣ/ , the pharyngeal fricative /h/ , the glottal fricatives /h/ , /h/.

- **Nasal:**
  Most Arabic consonants are produced orally with the velum raised. However, when the velum is lowered and the air stream is allowed to follow out through the nose, the resultant sound is called a nasal sound. Arabic nasal sounds are the bilabial nasal /m/ and the alveolar nasal /n/.
- lateral:
/I/ sound is an alveolar lateral consonant in Arabic. It is made by the front of the tongue pressing against the center of the alveolar ridge without contact with the sides of the hard palate, so the air stream escapes freely on the sides of the tongue.

- Trill:
It involves intermittent closure. This sound can be produced by tapping the tongue repeatedly against appoint of contact. In Arabic we have the alveolar trill/r/.

- Glide:
Glides are made without closure in the mouth. In Arabic, we have two glides; the bilabial /w/, and the palatal glide /j/.

<table>
<thead>
<tr>
<th></th>
<th>B</th>
<th>LD</th>
<th>I</th>
<th>A</th>
<th>D</th>
<th>P</th>
<th>V</th>
<th>G</th>
<th>Ph</th>
<th>U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop</td>
<td>ب</td>
<td>ت</td>
<td>ل</td>
<td>ح</td>
<td>ك</td>
<td>ق</td>
<td>غ</td>
<td>ع</td>
<td>ك</td>
<td>ق</td>
</tr>
<tr>
<td>Affricate</td>
<td>ح</td>
<td>ش</td>
<td>س</td>
<td>غ</td>
<td>خ</td>
<td>م</td>
<td>ن</td>
<td>ش</td>
<td>س</td>
<td>غ</td>
</tr>
<tr>
<td>Fricative</td>
<td>ف</td>
<td>ب</td>
<td>ث</td>
<td>ظ</td>
<td>س</td>
<td>ض</td>
<td>ك</td>
<td>ش</td>
<td>س</td>
<td>غ</td>
</tr>
<tr>
<td>Nasal</td>
<td>م</td>
<td>ن</td>
<td>م</td>
<td>ن</td>
<td>م</td>
<td>ن</td>
<td>م</td>
<td>ن</td>
<td>م</td>
<td>ن</td>
</tr>
<tr>
<td>Lateral</td>
<td>ل</td>
<td>ل</td>
<td>ل</td>
<td>ل</td>
<td>ل</td>
<td>ل</td>
<td>ل</td>
<td>ل</td>
<td>ل</td>
<td>ل</td>
</tr>
<tr>
<td>Trill</td>
<td>ر</td>
<td>ر</td>
<td>ر</td>
<td>ر</td>
<td>ر</td>
<td>ر</td>
<td>ر</td>
<td>ر</td>
<td>ر</td>
<td>ر</td>
</tr>
</tbody>
</table>

1.2.3 English and Arabic Sounds Contrasted

After describing the English and Arabic consonants in terms of their place and manner of articulation, it is evident that there are a number of differences between the sounds of the two languages. English has some sounds which are not available in Arabic, similarly, Arabic has a number of sounds that have no existence in English.

One of these differences is in the production of some consonants:

- /p/ does not occur in Arabic as a phoneme except in some loan or foreign words. Arab learners usually have difficulty with English contrasts such as /pæn/ pan and /bæn/ ban; /kæp/cap and /ka:b/ kab.
- /t/ and /d/ are dental in Arabic but alveolar in English. This difference is usually unnoticed by Arab learners of English, but it does not interfere with intelligibility.
- /v/ also does not occur in Arabic. Arab learners may have difficulty with English contrasts such as /væn/ van and /fæn/ fan.
- /θ/ and /ð/ do not usually occur in dialectal Arabic but they occur in standard Arabic. Instead /s/ and /z/ are used respectively. A common error is that some Arab learners of English use /s/ and /z/ for the English /θ/ and /ð/. Therefore, English consonants such as /pa: θ/path and /pa: s/ pass; /brɪθ/ breathe and /brɪz/ breeze are troublesome. Very often Arab learners make the same mistake when they learn standard Arabic. Therefore, the occurrence of /ð/ and /θ/ in standard Arabic and in English but not in dialectal Arabic causes a problem when pronouncing words with these two sounds.
There are some consonants that do not occur in standard Arabic. Such as /ʒ/, /ʧ/ and /ŋ/. Standard Arabic and some other dialects use /dʒ/ instead of /ʒ/ and some other dialects (e.g. Egyptian) use /g/. The choice of one form or another depends on the kind of dialect used. This causes some difficulty in the production of the English /ʒ/ which is often replaced by /dʒ/. Because the contrast between /dʒ/ and /ʒ/ in Arabic does not affect meaning (i.e. not phonemic), either one or the other is used by Arab learners in pronouncing English words having these two sounds. Usually /dʒ/ is wrongly used for /ʒ/. For example, words such as (occasion), (measure) and (explosion) are pronounced */əkeidʒən/, */medʒə/ and */iksplouʒən/ instead of /əkeiʒən/, /meʒə/ and /iksplouʒən/.

Moreover, /ʃ/ and /ʧ/ cause more problems as they are often confused especially in initial position. /ʃ/ is wrongly used for /ʧ/. For example (cheap) and (sheep) are pronounced */ʃip/ instead of /ʧip/ for the first and /ʃip/ for the second.

Also /ŋ/ does not occur in Arabic but it occurs as an allophone of /n/ before stop consonants such as /sIn/ sin and /sInŋ/ sing; /sInŋə/ sinner and /sInəŋ/ singer; /ræŋ/ ran and /ræŋ/ rang.

The clear and dark realizations of the sound /l/ occur in both English and Arabic but the dark /ɫ/ is restricted in Arabic to such words as /əɫɫa:h/. Therefore, a common mistake by Arab learners in pronouncing the lateral approximant /l/ is the excessive use of the clear /l/ in Arabic which is used in all positions: initial, medial and final, whereas in English the clear /l/ is used in initial and medial positions, and the dark /ɹ / is used in final positions and after consonants.

On the other hand, the sound /r/ is entirely different in the two languages. Arabic /r/ is a tongue tip trill, whereas English /r/ is a back alveolar. The initial problem is that Arab learners replace the Arabic /r/ for the English one. Secondly, English /r/ is replaced by a schwa /ə/ in final positions, e.g. /ridə/ reader but Arab learners tend to use /r/ in all cases, e.g. /ridær/reader.

Another difference between English and Arabic is consonant cluster. Consonant cluster is a group of consonants coming together without a vowel sound between them such as spin, skin, and steam. In English two, and three, four or even more consonants can follow each other to form consonant sequences or clusters, for example:

Two consonants: sphere, sleep, cuts, dogs.
Three consonants: street, risked, plunged, fix this.
Five consonants: mixed sweets.
Six consonants: first stream.
Seven consonants: tempts strangers.

Consonant clusters may come at initial or final positions in English words. The following are the commonest cases:

1. Two consonants at the beginning of words:
/s/ followed by /p, t, k, f, m, n, l, w, j/ as in spare, steam, sky, sphere, smile, sneer, slow, swear and suit.
/p/ followed by /r, w, j/ as in trim, twin tune.
/k/ followed by /l, r, w, j/ as in clip, crime, queen, queue.
/b/ followed by /l, r, j/ as in blame, bright, beautiful.
/d/ followed by /r, w, j/ as in dry, dwell, duty.
/g/ followed by /l, r/ as in glare, gross.
/fl/ followed by /l, r, j/ as in flare, frame, few.
/Θ/ followed by /r, w/ as in three, thwart.
/f/ followed by /r/ as in shrink.
/v, m, n, h/ followed by /j/ as in view, mute, news, huge.

2. Three consonants at the beginning of words:
/spr/ as in spray, spread, spring.
/str/ as in straggle, straw, stress, strict.
/skr/ as in scream, screen, screw, script.
/stj/ as in student, stupid.
/skj/ as in skew, (this sequence is rare in English).
/spl/ as in splash, splendid, split. /skw/ as in squad, square, squash, squeezes.

3. Two or more consonants at the end of the word:
 a. /s/ and /z/ at the end of nouns to form the plural as in facts, fields, books, dogs.
 b. /s/ and /z/ at the end of verbs in the third person in the present tense: eats, gives, risks, yields.
 c. /t/ and /d/ at the end of verbs in the past tense: as in wished, raised, risked, plunged.
 d. /Θ/ at the end of nouns and numbers: as in width, strength, seventh.
 f. /Θ/ followed by /s/ at the end of plural nouns and numerals as in breadths, sevenths.

4. Three or more consonants at word boundaries as in Best man, long skirt.

5. /n/ and /l/ are sometimes syllabic; that is, each of them occupies the place at the centre of the syllable which is usually occupied by a vowel. In this way, they create a consonant cluster. For example lesson is pronounced /lesən/ with a vowel /ə/ between /s/ and /n/, or /lesn/ without a vowel. /n/ means /θ/ that /n/ is syllabic. Similarly, the word level is pronounced /level/ with a vowel or /levl/ without a vowel. In this case, /l/ is syllabic.

The previous examples show how consonant clusters are available in English. They can occur at the beginning, medial, or at the end of word. However, this possibility is not available in Arabic, as consonants clusters in Arabic and English differ greatly.

As mentioned above, in English, two, three four or even more consonants can follow each other to form a cluster unit, e.g. sphere, street, next Sunday, etc., whereas in many forms of Arabic, there are no sequences of three or more consonants. The
following table summarizes English and Arabic consonant clusters or sequences (c= consonants):

<table>
<thead>
<tr>
<th></th>
<th>initial</th>
<th>medial</th>
<th>final</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>c</td>
<td>c</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>cc</td>
<td>cc</td>
<td>cc</td>
</tr>
<tr>
<td></td>
<td>ccc</td>
<td>ccc</td>
<td>ccc</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>cccc</td>
<td>cccc</td>
</tr>
<tr>
<td>Arabic</td>
<td>c</td>
<td>C</td>
<td>c</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>Cc</td>
<td>cc</td>
</tr>
</tbody>
</table>

As seen in this table, Arabic has no sequence of more than two consonants, which occur in medial and final positions but not in an initial position. English, on the other hand, has sequences of two or even more consonant clusters in all three positions. Furthermore, in connected speech the sequence may be even longer where one word may end with a consonant sequence and the next word with another. As a result, we can have sequences like /bæŋks klouzd/ banks closed.

This causes problems to Arab learners who often follow Arabic patterns of speech and add a vowel, which is called an intrusive vowel to break the consonant cluster. For example:

/sprIn/ spring becomes */sIrIn/
/grændfa: ða/ grandfather becomes /grændIfa:ða/
/bo: nt/ burnt becomes */bIrnt/ or /bernIt/

The first example shows an intrusive vowel at an initial position, the second example at a medial position, and the third example at a final position.

A similar breaking of the consonant cluster is found with the simple past tense of verbs ending in a voiceless consonant such as /wIʃt/ wished, /dropt/ dropped and /pa:st/ passed. These and similar verbs are often pronounced by Arab learners as */wIʃIt/, */droped/ and */pa: sed/ respectively.

The other difference between English and Arabic is represented in gemination. In Arabic a consonant may be doubled. This doubling of consonantal sounds in Arabic is accompanied by greater muscular tension. For example, the word ?adʒl which means “yes, indeed!” , or “certainly” or “appointed date” is pronounced with one /dʒ/ sound. However the word ?adʒdʒala, which means “postponed” is pronounced with two /dʒ/ sounds rather than one. This process is called gemination or consonant doubling. In English, there is no such doubling or gemination of consonant sounds, but in writing English; however, many letters are doubled, for example, allow, attack, beginning. Although such letters are doubled in writing, they are not geminated or doubled in pronunciation.

Because of these facts, i.e. gemination in Arabic pronunciation and doubling of letters in English orthography, Arabic learners of English tend to wrongly lengthen some English consonants. For example:

/dʒu:lri/ jewelry becomes /dʒu:llri/
/əlau/ allow becomes /əllau/
/ətæk/ attack becomes /ətæk/
So, Arab learners face some difficulty in pronouncing these words. The differences between English and Arabic in consonants and consonant clusters along with the traditional methods of teaching are assumed to be the cause of various errors students of both languages commit.

1.3 Vowels and diphthongs

1.3.1 English vowels

Todd (1987:14) defines the vowel sound as the sound that needs an open air passage in the mouth. In addition, vowels are produced with a relatively free flow of air. The quality of a vowel depends on the resonance space in the oral cavity and this is determined by the position of the mobile speech organs: the tongue, the lips, and the lower jaw. English vowels can be divided into two groups: short vowels and long vowels.

**Short vowels**

Short vowels are simple vowels, which are made without any stoppage of the air in the mouth. The short vowels of English have different characteristics. Consider:

- /ɪ/
  This short vowel is produced when the front of the tongue is in the half-close position, and the lips are spread. We find it in the word "hit".
- /ʊ/
  It is a short vowel. The central-back of the tongue is in half-close position and the lips are slightly rounded. It occurs in words like "book".
- /e/  
  When the front of the tongue is in the half-close position and the lips are unrounded the short vowel /e/ is produced. This appears in the words "egg, left".
- /ə/
  This sound is called schwa. It has a half close open central of the tongue position and lips are neutral. It appears in the words "about, banana". It represents the weak sound forms in English.
- /æ/
  This short vowel is formed with the front of the tongue in the half-open position and the lips are wide-open. This vowel is available in the word "attack".
- /ʌ/
  It is a central half-open short vowel, produced with lips in the neutral position. The words “run, uncle” are a case in point.

**Long vowels**

Long vowels are longer than others in the same position; they are written with "a length mark" /iː/. In English, there are five long vowels, as illustrated below.

- /iː/
  If the front of the tongue is in close position, lips are unrounded. This comes out with the long vowel / iː/ that occurs in the word "key, cheese, police".
/-uː/
The production of this sound involves the back of the tongue to be in the close position and the lips to be rounded but not lax. It is available in the words like, "food, rude, soup".

/-ɜː/
It’s a long vowel produced by the center of the tongue in half-close and half-open position with the lips in the unrounded position. It is available in words like "shirt, word, serve".

/-ɔː/
This long vowel is formed with the back of the tongue in the half-open position with rounded lips, as in," bought, law, all".

/-ɑː/
This vowel can be formed by the back of the tongue in the open position with unrounded lips and lowered jaw. It occurs in the words "far, clerk, part".

1.3.2 Arabic vowels

Vowels in Arabic are almost allophonic. Thus, they can be divided into two groups, which are short vowels and long vowels.

Short vowels
Unlike English vowels, Arabic vowels are represented by notations not by letters. Arabic have three short vowels as it appears below:

- /i:/
This vowel is called "kasrah" in Arabic. It appears under the alphabet in this shape (◌) as in, sir ﺮـﺳ "secret". The tongue in the half-close position forms /i/ vowel in Arabic. Here the lips are unrounded. This short vowel refers to the genitive case, which is assigned to the construct state in Arabic along with other nouns.

- /u:/
This is the Arabic "ḍammah", which is produced by the tongue in the half-close position with the rounded lips. In Arabic it is written as (◌) above the alphabet like the word "ṣum" ﺬـﺻ "deaf". It is used to express the nominative case, which is assigned to the subject and topic in Arabic.

- /a:/
This short vowel is represented by (◌) in Arabic and is called "fatha". The tongue in the half-open position and neutral lips form it. It appears in the word "Al walad-a", "the boy". It expresses the accusative case which is assigned to the object.

Long vowel
The long vowel is distinguished by the duration of time that it takes. Those long vowels are similar to short vowels in their position. Arabic has three long vowels. These are called huruufu l-madd ﺢـ۠وۡوـۡفـۡلـۡمـۡد "the letters of prolongation" (Hamad, 2003). The following points will give an idea about each one of such vowels.
- /ii/
This vowel is formed when we have the prolongation letter (y) ئ preceded by the notation ( ٌ) "kasrah'. For example, "fiil" ئل "elephant".

- /uu/
This long vowel is formed by the prolongation letter (w) ئ preceded by the notation ( ٌ) d, ammah. We find it in the word "ṭuul" طول "length".

- /aa/
The third Arabic long vowel is formed by the prolongation letter (a) preceded by the notation ( ٌ) fathah. It occurs in the word "qabala" قابل "met".

1.4. Diphthongs
1.4.1 English diphthongs

Diphthongs are a combination of two vowel sounds which take some duration of time as of a single long vowel. When diphthongs are produced, the tongue moves from one vocalic position to another. English has eight diphthongs. They are:

- Centering diphthongs
They end with a glide towards the central vowel /ə/:

- /ɪə/
The glide begins with a tongue position that is taken for /I/ and moves in the direction of /ə/. It is found in the words like "beer, fear".

- /ʊə/
It glides from a tongue position that is used for /ʊ/ toward the more open type /ə/. It appears in the words "sure, tour".

- /eə/
Its glide begins in the half – open front position and moves in the direction of more open variety of /ə/. For example, "chair, stare".

- Closing diphthongs
They end with a glide towards /i/ or towards /ʊ/:

- /eɪ/
The glide begins at a point behind the front open position and moves in direction of the position of /i/. We find it in the words "say, weigh".

- /ɔɪ/
The glide begins with tongue position that is for /ɔ/ and moves in direction of /i/, as in "toy, enjoy".

- /ai/
The glide begins at a point slightly behind the front open position, and moves in the direction of the position associated with /i/. It occurs in words such as "high, buy".

- /au/ 
Its glide begins at a central position of /ə/ and moves in the direction of the /ʊ/. It appears in the words "go, hello".

- /au/ 
The glide starts at a point between the back and front open position, and moves in the direction of the /ʊ/. As in "house, now, found".
**Arabic diphthongs**

A diphthong in Arabic is a combination of a vowel and consonant (Masloh, 1980). Arabic has two diphthongs. They are:

- `/aw/ "أو"`
  The glide begins from the vowels /a/ to the consonant /w/. As in /nawm/ "sleep".

- `/ay/ "أي"`
  The glide begins from the vowel /a/ to the consonant /j/. It is available in the word /bayt/ "house".

**English and Arabic Vowels Contrasted**

English and Arabic vowel systems highlight a number of differences:

- Firstly, it is found that vowels in English are more than Arabic vowels, that is, English work with phonemic vowel quality. On the other hand, Arabic vowels are more allophonic.

- Secondly, there are some English vowels that do not exist in Arabic, like /e, ə/.

- Thirdly, English vowels are affected by consonants following them, while Arabic vowels are not.

By looking at this contrast and how the two languages differ in their phonological systems, it is obvious that Arab learners will have difficulty in learning some vowels especially those that are not available in Arabic.

**1.5 Teaching English Phonology to Arab Learners**

**1.5.1 Collecting Data**

In this section, some classroom findings are listed to investigate the validity of the theoretical assumptions raised earlier. All data are gathered from a pronunciation class of 62 first-year students in the first academic year of 2011-2012. The class meets twice a week for a 90-minute session. Learners were not informed that their pronunciation is monitored for research purposes to minimize the Hawthorne Effect (Dornyei, 2007). The data collection took two stages: the first was a pre-training stage when a group of twenty students were asked to read a list of words and their pronunciation was monitored by the research assistant who took down notes. The second stage was a post-training exercise that came after 14 weeks of lecturing and training, and it involves the same twenty students. For ethical considerations, the informants were eventually informed that their responses were monitored for research purposes, and they expressed their oral consent and were assured that all names shall remain confidential and all data be used for research purposes only.

**1.5.2 Conducting the Study**

**1.5.2.1a Stage One: Pre-training Data**

The experiment, at this stage, includes asking the participants to individually read the list below. The list has the following words from Hancock (2003) which is the textbook used in this course:
To analyse the nominal quantitative data gathered at this stage, descriptive statistics (measures of frequency) is used, as Table 1 below shows:

<table>
<thead>
<tr>
<th></th>
<th>Frequency</th>
<th>Percent(%)</th>
<th>Valid Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. boat</td>
<td>7</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>2. cups</td>
<td>5</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>3. daughter</td>
<td>18</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>4. were</td>
<td>3</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>5. sung</td>
<td>10</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>6. cubs</td>
<td>20</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>7. shot</td>
<td>18</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>8. where</td>
<td>15</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>9. girl</td>
<td>10</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>10. scream</td>
<td>15</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>11. should</td>
<td>20</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>12. June</td>
<td>12</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>13. care</td>
<td>15</td>
<td>75</td>
<td>75</td>
</tr>
<tr>
<td>14. worked</td>
<td>9</td>
<td>45</td>
<td>45</td>
</tr>
<tr>
<td>15. about</td>
<td>8</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>16. child</td>
<td>7</td>
<td>35</td>
<td>35</td>
</tr>
<tr>
<td>17. case</td>
<td>20</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>18. heart</td>
<td>6</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>19. against</td>
<td>4</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>20. older</td>
<td>4</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>21. texts</td>
<td>12</td>
<td>60</td>
<td>60</td>
</tr>
</tbody>
</table>

Table 1. Frequency of Correct Pronunciation: Pre-training Stage

It seems evident that some of the listed words were of a higher difficulty if compared with the rest, and that explains the persistence of errors in some cases. By investigating such errors, it sounds clear that there are three reasons underlying these errors:

1. Relying on similar Arabic sounds;
2. Weak phonetic ability/low aptitude for oral mimicry;
3. Error fossilisation and/or faulty teaching.
First: Relying on Similar Arabic Sounds

In 1957, Lado put the foundations of what is called ‘Contrastive Analysis’:

We have here the major source of difficulty or ease in learning the foreign language…Those structures that are different will be difficult. (Lado1957, p.58)

This study opens with an endeavour to analyse the phonological systems in English and Arabic spotting areas of expected difficulty due to the obvious differences. However, this scheme should not be taken for granted. Not all what is different is difficult. For instance, the sounds /ɔɪ, ʊ, ɪ/ seem to be easy for the learners in the study group despite the fact that they do not exist in their first language L1. One possible explanation for this is that with such sounds, the glides make big, obvious movements in the resonance space that it is easy for the listener to grasp the quality of these sounds; unlike glides with shorter movements in the resonance space e.g. /eə, uə/. In the latter scenario, learners seem to rely on using similar Arabic sounds resulting in what is known as ‘inter-lingual errors caused by cross-linguistic influence.

Thus, in the list above, the words: boat, girl, sung, care, older, daughter, shot, should, about, and child are pronounced respectively in the following way:

*/bʊt/, */ɡeɪl/, */sʌngg/, */kɛr/, */uːdɑ/, */ʃʊt/, */ʃʊd/, */ʔbaut/, and */ʃaɪld/. Therefore, all the previous errors came as a result of relying on similar Arabic sounds.

Second: Weak phonetic ability/low aptitude for oral mimicry

Weak phonetic ability or what is also know as low aptitude for oral mimicry can be a main factor demonstrating why learners do not produce correct sounds (Kenworthy, 1987). This is specially true in case of sounds that are perceived as ‘very similar’. In the examples mentioned earlier, some learners cannot notice the difference between:

- /ɜː/ and /eə/ as in were, case, and where.
- /n/ and /ŋ/ as in sun and sung.
- /ʊ/ and /u/ as in should and shot.

And such low aptitude for oral mimicry resulted in faulty pronunciation where the learners assumed that the sounds mentioned above are ‘very similar’ or even ‘the same’.

Third: Error Fossilisation and/or Faulty Teaching

Learners might get used to a certain way of pronouncing a certain sound either because of faulty teaching or because they seem stuck in the journey between L1 and L2. Such ‘Interlanguage’ errors (Yule, 2009) might be the cause of a previous language teacher who mispronounced some sounds and is taken as a model for correct pronunciation. Alternatively, as the learner sets his inter-language journey, s/he might get stuck with fossilised errors and with little improvements to take record of.

In the examples to come, learners errors can be seen as a result of error fossilisation and faulty teaching:

- cups and cubs: Due to faulty teaching, learners assume that the final sound in the two words is the same. Most English teachers ignore or seem to overlook
the importance of teaching how sounds affect each other. That is, final voiceless consonants attract the voiceless allomorph \{s\}; whereas final voiced consonants attract the voiced allomorph \{z\}. As a result, \% 25 only managed to pronounce the word ‘cups’ correctly with the sound /s/ at the end, while all participants pronounced the word ‘cubs’ correctly. It seems evident that there is a tendency to pronounce the plural morpheme \{s\} as a /z/ sound. In later stages, it became evident to the research assistant that the learners never encountered a rule for this case, and with some students, applying the rule seemed to be difficult, and thus, regarded as ‘fossilised errors’.

- scream and worked: the two words have final consonant clusters which is a major area of difficulty for Arab learners as illustrated earlier in this paper. There is a tendency to break the cluster with \%75, and \%45 occurrence of correct pronunciation respectively. Prior to the training stage, the participants pronounced the words as /slkriːm/ and /wɔːkid/ and reported that this is how these words where taught to them.

As seen above, the reasons underlying pronunciation errors vary and sometimes overlap, as they can be justified by more than one factor. It is noteworthy to stay reminded that the frequency of errors detailed above was gathered prior any training. Will errors persist with training? That is what is dealt with in the coming section.

1.5.2.2a: Stage Two: Post-training Data

This stage, as pinpointed earlier, came after a fourteen-week training course on English pronunciation. The course presented many theoretical aspects similar to the earlier sections of this paper, i.e., students were taught the place and manner of articulation along with voicing qualities for English sounds, and later on, moved to do some contrastive analysis exercises by understanding the types of errors and the reasons behind such errors. To cut a long story short, the participants at this stage are aware of their errors and the possible explanations for such errors.

1.5.2.2b: Stage Two: Post-training Data Analysis

Similar to the first stage, the gathered nominal quantitative data are analysed using descriptive statistics (measures of frequency), as Table 2 below shows:

<table>
<thead>
<tr>
<th>Frequency of correct pronunciation</th>
<th>Frequency</th>
<th>Percent(%)</th>
<th>Valid Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. boat</td>
<td>17</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td>2. cups</td>
<td>17</td>
<td>85</td>
<td>85</td>
</tr>
<tr>
<td>3. daughter</td>
<td>20</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>4. were</td>
<td>12</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>5. sung</td>
<td>16</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>6. cubs</td>
<td>20</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>7. shot</td>
<td>18</td>
<td>90</td>
<td>90</td>
</tr>
</tbody>
</table>
Table 2. Frequency of Correct Pronunciation: Post-training Stage

Table 2 demonstrates a statistically significant difference in the frequency of correct pronunciation due to the fourteen-week training the informants received. For ease of comparison, Table 3, and Bar-Graph 1 below display the difference between the two stages:

<table>
<thead>
<tr>
<th>Item</th>
<th>Stage One</th>
<th>Stage Two</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. boat</td>
<td>35</td>
<td>85</td>
</tr>
<tr>
<td>2. cups</td>
<td>25</td>
<td>85</td>
</tr>
<tr>
<td>3. daughter</td>
<td>90</td>
<td>100</td>
</tr>
<tr>
<td>4. were</td>
<td>15</td>
<td>60</td>
</tr>
<tr>
<td>5. sung</td>
<td>50</td>
<td>80</td>
</tr>
<tr>
<td>6. cubs</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>7. shot</td>
<td>90</td>
<td>90</td>
</tr>
<tr>
<td>8. where</td>
<td>75</td>
<td>90</td>
</tr>
<tr>
<td>9. girl</td>
<td>50</td>
<td>75</td>
</tr>
<tr>
<td>10. scream</td>
<td>75</td>
<td>100</td>
</tr>
<tr>
<td>11. should</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>12. June</td>
<td>60</td>
<td>90</td>
</tr>
<tr>
<td>13. care</td>
<td>75</td>
<td>95</td>
</tr>
<tr>
<td>14. worked</td>
<td>45</td>
<td>85</td>
</tr>
<tr>
<td>15. about</td>
<td>40</td>
<td>80</td>
</tr>
<tr>
<td>16. child</td>
<td>35</td>
<td>75</td>
</tr>
<tr>
<td>17. case</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>18. heart</td>
<td>30</td>
<td>75</td>
</tr>
<tr>
<td>19. against</td>
<td>20</td>
<td>75</td>
</tr>
<tr>
<td>20. older</td>
<td>20</td>
<td>70</td>
</tr>
<tr>
<td>21. texts</td>
<td>60</td>
<td>80</td>
</tr>
</tbody>
</table>

Table 3: A Comparison between the Two Stages
Bar graph 1. A Statistical Comparison between the Two Stages

As seen in Table 3 and Bar Graph 1, stage two marks a significant progress in the participants’ pronunciation. With intensive training, the sounds: /əʊ/, /ɔː/, /ɜː/, /eər/, /dʒ/, /tʃ/ became more distinctive in the pronunciation of the participants. With the use of minimal pairs, recorded lists, drills, and theoretical definitions of sounds, the participants were more aware of the qualities of these sounds that appeared to be more problematic in the first stage.

Another tendency to notice here is that teaching the rules of final sounds attracting either voiced or voiceless morphemes came in handy with the words: cubs, cups, worked. Stage two witnesses a rise in correct pronunciation in this area.

More, the use of contrastive analysis to explain the differences between English and Arabic minimised the occurrences of breaking the consonant clusters. This is obvious if percentages are traced with the words: scream, worked, against, and texts.

Overall, it seems evident that long training on practical and theoretical aspects of pronunciation can be fruitful. Results are not expected to be a hundred percent change because learners, even when aware of the rules, tend to over-generalise or fossilise some words. Changing an old pronunciation habit takes time and requires patience and perseverance on the language tutor’s behalf.

1.6 Recommendations and Pedagogical Implications

Indeed, the phonological systems of English and Arabic have a lot of asymmetries. Therefore, there are lots of areas of confusion and mispronunciation. As a result, it is high time for language teachers to intervene with suitable solutions. For example, teachers need to use various teaching activities which sensitise learners to differences in the sound systems of Arabic and English. Introducing some theoretical elements could make the picture clearer. i.e. with adult learners, it could be a good idea to introduce the place and manner of articulation for consonants and the cardinal vowel scheme for vowels. It also helps when teachers pinpoint some of the aspects of contrastive analysis to raise awareness regarding inter-lingual errors. In addition, playing with minimal pairs is a good way to strengthen the phonetic ability of the
learners and to sharpen their ears to pick the subtle differences between what seems
the ‘same’ to them. Most importantly, constant exposure to English sounds guarantees
a better level of performance. Teachers need to welcome diversity and technology.
Students will be excited to hear someone else’s voice in the classroom. Introducing
recorded materials and audio-visual aids is not only motivating, but also rewarding. It
is noteworthy to say that a lot of teachers worry about ‘overburdening’ their students
with a lot of theoretical definitions, terms, and rules lest they might complicate the
picture. However, this research suggests that theoretical awareness will not only
improve the learners’ pronunciation, but will justify learners’ errors, as well. Last, a
golden rule for teaching pronunciation, or any other language aspect, is change in the
linguistic performance of the learner is not likely to happen soon. Learners could
grasp the rules, the theories, etc, but this will not lead to a direct change in their
linguistic performance, rather, it enriches their linguistic competence that could take
time to be transferred into performance, in Chomskyian terms.
References


