Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327
مجلة اللغة الانكليزية في العالم العربي

AWEJ Volume.5 Number.4, 2014

Pp.365-380

Towards Processing Arabic Minimal Syllable Automatically

Mohammed Dib
University of Mascara, Algeria

Abstract

The purpose of this paper is to try to treat the Arabic minimal syllable automatically, so as to use Arabic in the field of artificial intelligence. To this effect three technological tools are used; *Gold wave ,SFS (Speech Filing System), and Neural Net Works* to recognize automatically the minimal syllable located in first, mid, and final position of three Arabic words recorded by forty Algerian speakers of different age and sex. Eight experiments have been done in this work where the sounds have been recorded in Gold wave and treated in SFS and trained in NNW. The result show that The optimal neural net work is that of *non- ordered data* with one layer, five nodes and 150 steps because it has given an error rate of 0.0032.The findings suggest the application of this type of neural net works in all syllables and all languages too because the same principle can be used in all languages.

Key words: Arabic, minimal, syllable, Gold wave, SFS, Neural Net Works

Introduction

It is unreasonable to ignore the importance of rapid changes resulting from globalization whose octopus hand tries to cover both of micro and macro fields and technological process in particular, therefore should accept and cope with these changes so as to serve Arabic language, its community, and to solve problems such as human-machine communication. This paper treats how to process Arabic minimal syllable automatically, and there is no doubt that the motives behind this work are the sharp deficiencies of the Arabic language as far as computer sciences are concerned. , and the problems that Arabic language faces nowadays with regard to globalization which tries to put it back in the black despite its large number of words, its users and its system in terms of flexibility and exactness. So English became the turbine engine of globalization, while Arabic loses its share in some Arabic countries and is even neglected by its speakers and researchers as well who do not endeavour to develop it as English researchers do. So, is this neglection due to its complicated system, or to other different reasons?

In this research we would justify that:

- Arabic is a flexible language
- Its limited number of vowels can help easily in automatic speech recognition

The main purpose of this experiment is to treat the Arabic minimal syllable automatically, so as to use Arabic in the field of artificial intelligence. In this account we try to identify limitations that point the way to this research relying upon the aforementioned technological tools as means of recognition.

A brief sketch of the outline can show two parts; a theoretical side and practical one. The theoretical side deals with speech recognition, intelligence, and neural net works. The practical side, however, contains syllable, types of syllables in Arabic, corpus, informants, syllable in question, automatic treatment, result and a conclusion.

1. **Speech recognition**

Speech recognition can be viewed as a communication problem between man and machine, that is, machine tries to recognize a word sequence pronounced by a speaker whose speech production process is very complicated and goes through some stages; the brain first generates the text which is composed of word strings, then goes to the acoustic process where converted into audible wave form (WuChou,A ,2009p.14) In other words, an analog signal is converted into an acoustic signal or digital one to get classified decisions, or a variable stimulus is transformed into a constant response. (Plomp.R, 2002 p.94)

Speech recognition is an inverse operation which starts from the speech wave form and ends in decoding a message. It is a mechanism able to decipher speech signal coming from the vocal tract or nasal cavity represented in a sequence of linguistic units found in the message that the speaker wants to transmit (Rabiner.L & Juang.B.H, 1993p.4).

The final goal of automatic speech recognition is communication between man and machine. The interaction has known many applications due to the rapid growth of devices and technological programmes (Peinado.A. M,2002p.1). Automatic speech recognition requires knowledge of many fields such as signal processing, acoustic phonetics, patterns recognition, communication, theory of information, and physiology.

2. **Intelligence**

No exact definition is found for the word intelligence , however some attempts were made to clarify the term. It is the ability to reason in relation to the solution of a new

problem(Gutman.R.D,2009p.55).There are a lot of types of intelligence which can be summarized into seven forms of knowledge; linguistic intelligence, mathematical and logical intelligence, spatial intelligence, musical intelligence, kinesthetic bodily intelligence, personal intelligence and interpersonal intelligence.(Brown.D.H,2007p.100)

As for artificial intelligence, it can be said that it is the reproduction of all types of human intelligence in computational programmes (The Hutchinson Encyclopedia,1999).The goal of artificial intelligence is to simulate human intelligence. The idea came after having investigated tremendous fields such as medical field and noticed the work of the human body and the brain in particular, and understood the process of language therefore dealt with natural language processing.

2.1 Neural Net works

It is a challenge nowadays to understand the human brain's work. It is undoubtedly that the best way which can enable us to investigate data processing in the human brain mathematically and computationally is the modelisation of the neural net works. They can be defined as a mathematical system which contains processors similar to the brain's cells. They contain a set of nodes that gather input from different sources then send them to other nodes which in their turn resend them to other nodes. They can get very complex input and represent them in a very simple output. They contain three layers; input layer, output layer, and a hidden layer in between, where each processor is in contact with another one throughout synapses (The Hutchinson Encyclopedia, 1999). They represent many mathematical models of human brain's functions such as: comprehension, calculation and memorization. There are many different types of neural net works, however they have same four basic attributes summarized as: a set of processing units, a set of communication, a computing procedure, and a training procedure (Tebelskis.J, 1995 p.28).

3. Syllable

No definition is completely satisfactory for the term syllable, but any attempt at a definition should take into consideration that it is a prominence peak surrounded by a cluster of consonants, however sometimes syllable boundaries are put aside and the question whether some peaks such as / s/ in stop or in other words are not considered as syllables are to be avoided (San.D, 2008 p.36) . Another definition states that syllable is related to chest pulse but does not refer to syllable boundaries, more simply put Gimson notes that the double chest pulse does not seem clearly in the word *seeing* [**si:ɪŋ]**and the pulse theory cannot decide whether the word *beer* [**bɪər]** contains two syllables in American pronunciation. This question generates some doubts on whether syllables are linguistic units or not. Chomsky, Halle, Steriade, Gimson, Belvins do not consider them as phonological units (Duanmu.S,2008) Despite all this, syllable appears clearly in some cases; for example people agree that the word Canada entails three syllables, in contrast the word America comprises four syllables (.Tebelskis.J,1995)

3.1Syllable in Arabic

Syllable in Arabic always starts with a consonant and ends either with a vowel, called the open syllable, or with a consonant, called the close syllable. This means that the word in Arabic never starts with a consonant cluster, whatever the manner of consonant is. The following example is a good demonstration of the point. (***uktub***) is the imperative form of the verb "to write" and is impossible to say (***ktub***) because the Arabic phonological system rejects consonant cluster, so we

brought the ***hamza*** which stands for the vowel /u/. The same thing can be said for the Greek word (**platoon**) which has become (***Aflaton*****)** in Arabic and the French word (**franc)** which has changed into (***Ifrange***) (Henry.F, 2007p.43)

Generally speaking, the syllable is an association between a consonant/s and a vowel/s. Roman Jakobson defines the syllable as a group of structure which encompasses two associated phonemes whose degree of aperture is different; one of a smaller degree and the other of a bigger one. (Abdelouhab.H,1984p.27). Amber Crombie, however, argues that speech hinges upon breathing, and air exhaling is similar to pulsation; each muscle contraction resulting from the air pressure forms a chest pulse, and each chest pulse in its turn forms a syllable. The pattern of chest pulses is the basis of human language. (Mubark.H, 1997p.65)

A number of source said that the dividing speech into syllables goes back to the period where Arabic language was an oral language and relied only on the listening process to transfer literature and arts (Rumani, 1976p.60). Aljahid, one of the Arab grammarians, used the term syllabification which means segmentation of speech. He stated that the sound is a device of speech whose role is syllabification and connection (Aljahid, 2002p.79).

3.2 Structure of Arabic syllable

Syllable is a combination of a consonant and a vowel which goes on a par with the system of each language in forming its syllable structure. It relies upon the respiratory rhythm. The minimal syllable in Arabic is formed by one consonant followed by a short or a long vowel; this means that a sequence of two consecutive consonants is unacceptable except in case of pause.

The syllable in Arabic never starts with two consonants or a vowel, this is why syllables in all languages consist of vowels as centers preceded or followed by consonants in spite of the differences that exist between languages over the location of consonants, but in some cases syllables may be formed without a vowel .The Czech words are a good demonstration of the point:(***prno***, ***vltava***) where these syllables consist of consonants only.

One third of the studied languages use consonants only to form syllables (Claude.H, 1997p.24) German and English are among those languages but Arabic is excluded from that since its phonological system rejects two consecutive consonants. The German word (***abend***) is pronounced /***abant***/ in careful speech, where as in connected speech is pronounced (***abnt***) or (***abmt***), likewise English words **bottle** and **button** are pronounced /**bɒtl**/,/**bʌtn**/with the omission of schwa vowel and sometimes the syllabic consonant appears as a result. (Peter.R,1991 p.106)

3.3 Types of syllables

According to (Kaddour.A.M,1999), there are five types of syllables which are as follows: minimal syllable, closed long syllable, open long syllable, long syllable closed with a consonant, and long syllable closed with two consonants.

3.3.1 Minimal syllable, it consists of a consonant and a vowel. It is a meaningful linguistic unit, and consists in prepositions e.g. bi (by),fi (in),li (for)…etc.

3.3.2Closed long syllable, it consists of a consonant, a vowel and a consonant.eg. *mithl* (like) min (from) bal(rather)

3.3.3Open long syllable, it consists of a consonant and two vowels .e.g. *maa and haa*

Two forms related to the pause.

3.3.4Long syllable closed with a consonant, it is made up of a consonant, two vowels, and a consonant e.g. kaan (was)

3.3.5 Long syllable closed with two consonants, it is composed of a consonant, a vowel, and two consonants e.g. *karb, fadl*.

4. Data base

Arabic language contains 28 consonants, and 6 vowels; 3 long vowels represented in three consonants *waw,ja and alif*, and 3 short vowels represented in three diacritic marks *fathat, dammat* and *kasrat*,.The list of consonants starts with the /ʔ/ sound known as ***hamza*** and ends with the /j/ sound. To get a simple unit one short vowel is added to a consonant e.g. (b+a,), by contrast to get a complex unit many combinations are made as shown above in types of syllables. It should be noted that the data is tabulated below. As for the words selected for study, they are three words containing the minimal syllable/ **ʔa/** located in first, mid and final position and were recorded by 40 speakers then segmented and as a result 120 syllables were obtained. Table 1 and table 2 may show this clearly

5. Syllable's Recognition Using Neural Net Works

In order to process sound, three fundamental phases should be taken into account; Pre-Treatment, Treatment and After-Treatment phases. In the Pre-Treatment phase, words are recorded (known as speech prior analysis) over a microphone, and then segmented into syllables to process the sound. It is the starting point which helps in automatic treatment. Then move to the next phase known as treatment phase where an analog signal is converted into a digital one in order to get the spectrum known as the identity card of the sound because it contains numerical values, that is, frequencies, amplitude, and periodicity. After that, we move to the final phase called after-treatment phase where characteristics of the signal are injected in the neural system as an input e.g. formant1 ,formant2,formant3,formant, periodicity, and amplitude, then try to train the system to make an association between input and output.

It should be noted that the process of recognition requires some symbols to codify data so as to be comprehended by the machine using neural net works. The code of this data is represented into letters and numbers. The letters used for code are: A,h,d,m ,g and f .A represents the minimal syllable known as ***hamza*** in Arabic and a short vowel**,** the **(h)** stands for (*homme*), the (**d)** means (*debut*),that is, the beginning, the(**m)** means the medial syllable ,and the **(g)** means garçon. While the letter f has two representations; one means final and the other means femme (woman). As for the numbers, they represent speakers. Tables 3 shows the words that were recorded by 40 informants of different sex and age; 10men, 10women, 10boys, and 10girls .Then segmented to have the minimal syllable referred to as ***hamza*** and the short vowel(a) located in first, mid and final position.

As indicated in Table 4 and 5, numerical values concerning referential and test samples are displayed for the process of recognition. We mean by referential sample input and output of minimal syllable located in first, mid and final position. Input encompasses numerical values of formants, amplitude and period, while output represents the recorded sounds**.** Training neural net works to recognize sounds requires three things ;frequency of formants, amplitude and period which are put in the upper layer as nodes, then connected to a hidden layer which is connected to a lower layer that represents output.

Table*1. *Vowels

Number	***Arabic written form***	***Latin written form***	***phonetic transcription***
short vowels			
1	ُ	***DAMMA***	**u**
2	َ	***FATHA***	**a**
3	ِ	***KASRA***	**I**
Long vowels			
1	و	***WAW***	**u:**
2	ا	***ALIF***	**a:**
3	ي	***YAE***	**i:**

Table1 shows the types of vowels in Arabic, number of vowels, Arabic written form, Latin written form, and phonetic transcription as well. The first part of the table concerns the short vowels, while the second part regards the long vowels. The number of short vowels is three which are: *DAMMA, FATHA, KASRA,* and stand for the sound [u], [a], and [I] in English.

Table 2. ***Minimal syllables***

Number	***Arabic written form***	***Latin written form***	***Phonetic transcription***
1	ءَ	a	/ʔa/
2	ءُ	u	/ʔu/
3	ءِ	ɪ	/ʔɪ/

Table 2 displays three minimal syllables that contain a consonant known as ***hamza*** and a vowel represented in diacritic marks (ًٌٍ).The three syllables are regarded as short vowels in English (a, u, ɪ), but in Arabic as glottal stop and a vowel.

Table 3. ***Syllables***

Word	***Phonetic transcription***	***First syllable***	***Medial syllable***	***Final syllable***
أمر	/ʔamara /	أ /ʔa/		
ثأر	/θaʔara/		أ /ʔa/	
قرأ	/qaraʔa/			أ /ʔa/

Table3 shows the three recorded words in Arabic,***أمر*** , ***ثأر***, and ***قرأ*** , that contain the minimal syllable /ʔa/ in first, mid, and final position ,and their phonetic transcription as well. They are written in Arabic and transcribed according to the International Phonetic Association.

Table4.*Referential samples*

Sound	**F1**	**F2**	**F3**	**F4**	**Amp**	**Per**
Ahd1	**658Hz**	**1076Hz**	**2668Hz**	**0Hz**	**0.1566dB**	**0.0001667**
Ahm1	**641Hz**	**1037Hz**	**1692Hz**	**2504Hz**	**0.2608dB**	**0.0001667**
Ahf1	**580Hz**	**928Hz**	**1621Hz**	**2633Hz**	**0.2602dB**	**0.0001667**
Ahd2	**775Hz**	**1882Hz**	**2152Hz**	**0Hz**	**0.1835dB**	**0.0001667**
Ahm2	**768hz**	**1202Hz**	**2231Hz**	**0Hz**	**0.0946dB**	**0.0001667**
Ahf2	**700Hz**	**1115Hz**	**2144Hz**	**0Hz**	**0.2168dB**	**0.0001667**
Ahd3	**683Hz**	**1038Hz**	**1359Hz**	**2269H z**	**0.1609dB**	**0.0001667**
Ahm3	**721Hz**	**1195**	**2480**	**0Hz**	**0.1600dB**	**0.0001667**
Ahf3	**801Hz**	**1098Hz**	**1986Hz**	**0Hz**	**0.2894dB**	**0.0001667**

Ahd4	**659Hz**	**1154Hz**	**2445Hz**	**0Hz**	**0.1632dB**	**0.0001667**
Ahm4	**625Hz**	**1025Hz**	**1649Hz**	**2489Hz**	**0.1912dB**	**0.0001667**
Ahf4	**640Hz**	**1069Hz**	**2660Hz**	**0Hz**	**0.1635dB**	**0.0001667**
Ahd5	**670Hz**	**1140Hz**	**2714Hz**	**0Hz**	**0.1306dB**	**0.0001667**
Ahm5	**698Hz**	**1204Hz**	**2640Hz**	**0Hz**	**0.1632dB**	**0.0001667**
Ahf5	**652Hz**	**1014Hz**	**2696Hz**	**0Hz**	**0.1628dB**	**0.0001667**
Afd1	**775Hz**	**1182Hz**	**2152Hz**	**0Hz**	**0.1840dB**	**0.0001667**
Afm1	**768Hz**	**1202Hz**	**2231Hz**	**0Hz**	**0.0949dB**	**0.0001667**
Aff1	**700Hz**	**1115Hz**	**2144Hz**	**0Hz**	**0.2194dB**	**0.0001667**
Afd2	**803Hz**	**999Hz**	**1649Hz**	**2541Hz**	**0.1245dB**	**0.0001667**
Afm2	**823Hz**	**1087Hz**	**1536Hz**	**2549Hz**	**0.1251dB**	**0.0001667**
Aff2	**898Hz**	**1387Hz**	**2340Hz**	**0Hz**	**0.1274dB**	**0.0001667**
Afd3	**792Hz**	**1132Hz**	**1534Hz**	**0Hz**	**0.0972dB**	**0.0001667**
Afm3	**896Hz**	**1525Hz**	**0Hz**	**0Hz**	**0.0972dB**	**0.0001667**
Aff3	**835Hz**	**1399Hz**	**0Hz**	**0Hz**	**0.1603dB**	**0.0001667**
Afd4	**717Hz**	**1058Hz**	**1454Hz**	**2520Hz**	**0.2603dB**	**0.0001667**
Afm4	**835Hz**	**1020Hz**	**1510Hz**	**2644Hz**	**0.1312dB**	**0.0001667**
Aff4	**796Hz**	**1065Hz**	**1339Hz**	**2645Hz**	**0.1953dB**	**0.0001667**
Afd5	**875Hz**	**1462Hz**	**2625Hz**	**0Hz**	**0.1962dB**	**0.0001667**
Afm5	**813Hz**	**1002Hz**	**1520Hz**	**2638Hz**	**0.1286dB**	**0.0001667**
bAff5	**871Hz**	**1419Hz**	**2571Hz**	**0Hz**	**0.2577dB**	**0.0001667**
Agd1	**683Hz**	**1038Hz**	**1359Hz**	**2269Hz**	**0.1600dB**	**0.0001667**
Agm1	**721Zh**	**1165Hz**	**2480Hz**	**0Hz**	**0.1603dB**	**0.0001667**
Agf1	**801Hz**	**1098Hz**	**1986Hz**	**0Hz**	**0.2899dB**	**0.0001667**

Agd2	**911H**	**1257Hz**	**2225Hz**	**0Hz**	**0.1563dB**	**0.0001667**
Agm2	**564Hz**	**1550Hz**	**2520Hz**	**0Hz**	**0.1872dB**	**0.0001667**
Agf2	**646Hz**	**918Hz**	**1255Hz**	**2180Hz**	**0.2214dB**	**0.0001667**
Agd3	**537Hz**	**1127Hz**	**1654Hz**	**2051Hz**	**0.1878dB**	**0.0001667**
Agm3	**541Hz**	**1032Hz**	**1658hz**	**1818Hz**	**0.1560dB**	**0.0001667**
Agf3	**653Hz**	**983Hz**	**1352Hz**	**2299Hz**	**0.2275dB**	**0.0001667**
Agd4	**853Hz**	**1096Hz**	**1584Hz**	**2411Hz**	**0.1953dB**	**0.0001667**
Agm4	**763Hz**	**1243Hz**	**1635Hz**	**2528Hz**	**0.1291dB**	**0.0001667**
Agf4	**785Hz**	**1215hz**	**1782Hz**	**2416Hz**	**0.2539dB**	**0.0001667**
Agd5	**688Hz**	**972Hz**	**1347Hz**	**2631Hz**	**0.1918dB**	**0.0001667**
Agm5	**660Hz**	**903Hz**	**1460Hz**	**2595Hz**	**0.1600dB**	**0.0001667**
Agf5	**743Hz**	**905Hz**	**1271Hz**	**2605Hz**	**0.1606dB**	**0.0001667**

Table 4 deals with the characteristics of the sound such as formant1, formant2, formant3 formant4, period, and amplitude which are all injected in neural net works during the training phase except for period which is the same for all the sound, therefore is discarded. It worth mentioning that there are five formants with regard to speakers, however formant 5 is missing because there are no nasal neighbouring sounds in the aforementioned Arabic words. Formants are places of resonance that are classified from low to high and are larynx, pharynx, back part of the mouth, front part of the mouth, and nasals, and amplitude is the maximum movement away from the place of rest. (Connor.J.D.O,1973).The table also shows numerical values that concern the formants of the first group of speakers which consist of 20 participants and were asked to pronounce the syllable /ʔa/which is regarded as a sound in acoustic phonetics. The numerical values mentioned in table 4 concern each participant pronouncing /ʔa/ in first ,mid and final position of the word i.e., each speakers utters the syllable three times. This can be displayed according to the following symbols: A,h,d,m,g and f ,where A represents the minimal syllable, **(h)** represents adult speakers (*homme*), (**d)** means (*debut*),that is, the beginning, (**m)** means syllable located in mid position, and **(g)** means garcon, where there are two groups; boys and girls. As for the letter f, it is used to represent two things; final position and female adult speaker (femme). Finally numbers,1,2,3,4,5 are used to represent speakers.

Table5. ***Test samples***

Sound	F1	F2	F3	F4	Amp	Per
Ahd6	796Hz	1239Hz	2206Hz	0Hz	0.2603dB	0.0001667
Ahm6	733Hz	1213Hz	2627Hz	0Hz	0.1320dB	0.0001667
Ahf6	747Hz	1143Hz	2684Hz	0Hz	0.2278dB	0.0001667
Ahd7	714Hz	1075Hz	2028H	0Hz	0.1638dB	0.0001667
Ahm7	672Hz	1085Hz	2054Hz	0Hz	0.1652dB	0.0001667
Ahf7	565Hz	1552Hz	2505Hz	0Hz	0.2275dB	0.0001667
Ahd8	663Hz	1134Hz	2326Hz	0Hz	0.1641dB	0.0001667
Ahm8	602Hz	1037Hz	2374Hz	0Hz	0.1353dB	0.0001667
Ahf8	602Hz	1037Hz	2374Hz	0Hz	0.1359dB	0.0001667
Ahd9	670Hz	1068Hz	1750Hz	2641Hz	0.1600dB	0.0001667
Ahm9	661Hz	1088Hz	2543Hz	0Hz	0.0998dB	0.0001667
Ahf9	632Hz	976Hz	1689Hz	2634Hz	0.1312dB	0.0001667
Ahd10	714Hz	1209Hz	2474Hz	0Hz	0.1635dB	0.0001667
Ahm10	691Hz	1167Hz	2329Hz	0Hz	0.1910dB	0.0001667
Ahf10	642Hz	994Hz	1498Hz	2315Hz	0.1626dB	0.0001667
Afd6	466Hz	1210Hz	1659Hz	0Hz	0.1635dB	0.0001667
Afm6	502Hz	1217Hz	1681Hz	0Hz	0.2280dB	0.0001667
Aff6	661Hz	1030Hz	1580Hz	2477Hz	0.1970dB	0.0001667
Afd7	777Hz	1082Hz	1604Hz	2528Hz	0.0966dB	0.0001667
Afm7	946Hz	1581Hz	2692Hz	0Hz	0.1280dB	0.0001667
Aff7	687Hz	969Hz	1491Hz	2562Hz	0.1009dB	0.0001667
Afd8	691 Hz	971Hz	1387Hz	2475Hz	0.1280dB	0.0001667
Afm8	599Hz	880Hz	1448Hz	2815Hz	0.1632dB	0.0001667
Aff8	505Hz	1009Hz	1260Hz	2522Hz	0.1973dB	0.0001667

Afd9	**815Hz**	**1053Hz**	**1540Hz**	**2601Hz**	**0.1956dB**	**0.0001667**
Afm9	**849Hz**	**1384Hz**	**2610Hz**	**0Hz**	**0.1312dB**	**0.0001667**
Aff9	**858Hz**	**1428Hz**	**2706Hz**	**0Hz**	**0.1629dB**	**0.0001667**
Afd10	**1020Hz**	**1461Hz**	**2366Hz**	**0Hz**	**0.1944dB**	**0.0001667**
Afm10	**1041Hz**	**1599Hz**	**0Hz**	**0Hz**	**0.1312dB**	**0.0001667**
Aff10	**752Hz**	**888Hz**	**1308Hz**	**2351Hz**	**0.2278dB**	**0.0001667**
Agf6	**791Hz**	**983Hz**	**1353Hz**	**2274Hz**	**0.2545dB**	**0.0001667**
Agd7	**795Hz**	**1184Hz**	**1601Hz**	**0Hz**	**0.2231dB**	**0.0001667**
Agm7	**565Hz**	**1552Hz**	**2505Hz**	**0Hz**	**0.2228dB**	**0.0001667**
Agf7	**634Hz**	**1015Hz**	**1547Hz**	**2295Hz**	**0.2545dB**	**0.0001667**
Agd8	**873Hz**	**1074Hz**	**1497Hz**	**2377Hz**	**0.1289dB**	**0.0001667**
Agm8	**896Hz**	**1240Hz**	**1595Hz**	**2428Hz**	**0.1956dB**	**0.0001667**
Agf8	**801Hz**	**1028Hz**	**1235Hz**	**2269Hz**	**0.1312dB**	**0.0001667**
Agd9	**565Hz**	**1552Hz**	**2505Hz**	**0Hz**	**0.1635dB**	**0.0001667**
Agm9	**925Hz**	**1538Hz**	**2373Hz**	**0Hz**	**0.1317dB**	**0.0001667**
Agf9	**776Hz**	**1205Hz**	**1979Hz**	**0Hz**	**0.1915dB**	**0.0001667**
Agd10	**947Hz**	**1318Hz**	**2645Hz**	**0Hz**	**0.2205dB**	**0.0001667**
Agm10	**918Hz**	**1434Hz**	**2684Hz**	**0Hz**	**0.1439dB**	**0.0001667**
Agf10	**581Hz**	**887Hz**	**1191Hz**	**2513Hz**	**0.1919Bd**	**0.0001667**

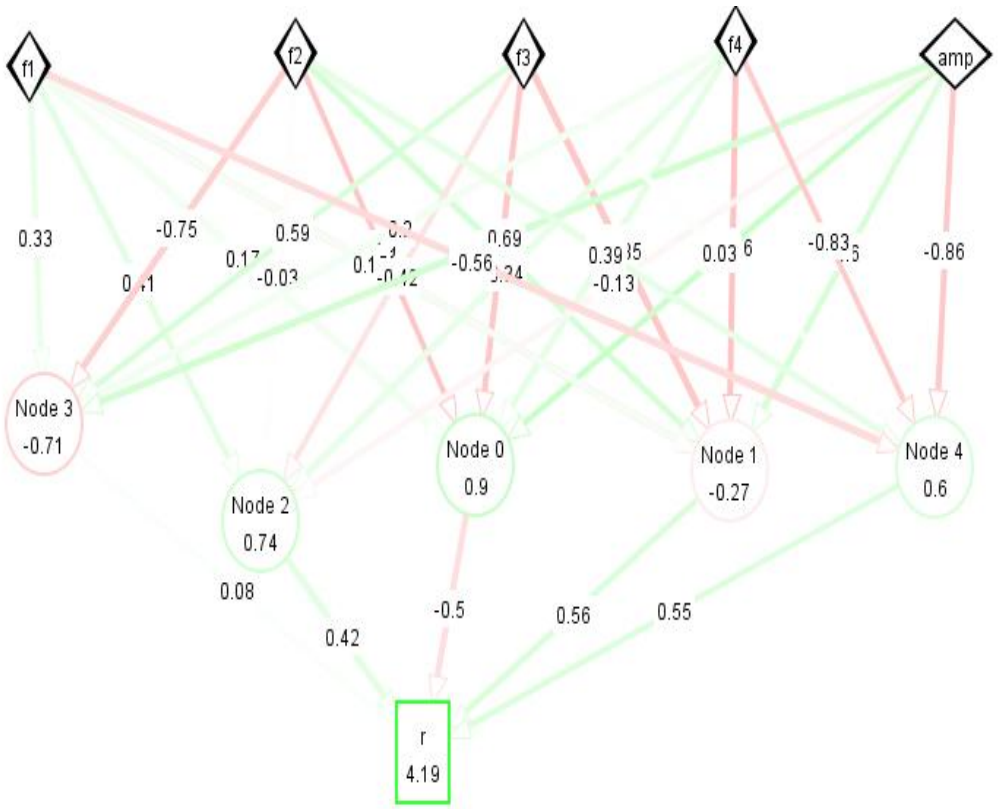
Table 5 is concerned with 20 other participants pronouncing the syllable of the Arabic words mentioned above. It shows numerical values of four formants and amplitude of the sound (syllable). Symbols are also used to represent the syllable; located in first, mid, and final position, and are similar to those shown in previous chart, however numbers representing speakers are different because they represent another group of speakers which consists of 5men ,5women,five boys, and five girls.

Table6. ***Types of Neural net works***

NNW with ordered data				
Number of NNW	**Number of hidden layers**	**Number of nodes in hidden layer**	**Number of steps**	**Error rate**
1	**1**	**2**	**150**	**0.9783**
2	**1**	**3**	**150**	**0.9703**
3	**1**	**4**	**150**	**0.9791**
4	**1**	**5**	**150**	**0.9791**
NNW with non- ordered data				
Number of NNW	**Number of hidden layers**	**Number of nodes in hidden layer**	**Number of steps**	**Error rate**
5	**1**	**2**	**150**	**0.0074**
6	**1**	**3**	**150**	**0.0060**
7	**1**	**4**	**150**	**0.0047**
8	**1**	**5**	**150**	**0.0032**

Table 6 displays the eight types of NNW used in our experiment; four types with ordered data, while other four with non-ordered data. The left column displays numbers of the NNW which are eight. The column where number 1 is repeated eight times indicates that the eight NNW contain one hidden layer, whereas column which contains numbers 2, 3, 4 and 5 concerns the number of nodes of the hidden layer. The column where 150 appears eight times stands for the number of steps used in each attempt to get the result. Finally the last column shows the error rates of each experiment.

Figure1. ***Design of Neural Net Works***



150 Steps .Error rate 0.0032

Figure 1 represents the design of the neural net work that consists of three layers; an input layer, an output layer and a hidden layer in between. The input layer contains five nodes where the first four nodes represent the place of resonance of the sounds known as formants in acoustic phonetics, while the last node represents the amplitude of the sound. The output layer contains only one node that concerns the result. As for the layer in between, it comprises five nodes called processors that deal with mathematical operations. It should be noted that the five nodes of the input layer are constant .i.e., the input layer comprises five characteristics only; formant1 (F1) formant2 (F2) formant3 (F3) formant4 (F4), and amplitude (Amp), however numerical values of these characteristics (formants and amplitudes) are variable, because data base contains 40 speakers and three syllables i.e., 120 units (syllables) of five characteristics each, therefore 600 characteristics are injected in the programme of Neural Net work.

Data Analysis

Training the neural net works for syllable recognition have gone through many experiments represented in eight NNW as shown in table 6; four neural net works with ordered data, while four others with non ordered data. The reason behind these experiments is to demonstrate which one can bring better results; As for NNW with ordered data, they encompass neural net work number 1 which consists of one hidden layer, two nodes, 150 steps and the error rate is 0.9783, neural net work number 2 of one hidden layer of three nodes150 steps and error rate of 0.9703. Neural network number 3 of hidden layer of four nodes ,150 steps and error rate 0.9791.Net work number four , however, of one hidden layer of five nodes ,150 steps and error rate of 0.9791.

As regard neural net works with non ordered data, they contain NNW number 5 which comprises one hidden layer of two nodes, 150 steps and error rate 0.0074, NNW6 with one hidden layer of three nodes, 150 steps and error rate0.0060.NNW number 7 with one hidden layer of four nodes, 150 steps and error rate of 0.0047, and NNW number 8 with an error rate of 0.0032, one hidden layer five nodes and 150 steps. The optimal NNW is net work number 8 with one hidden layer five nodes and 150 steps and error rate of 0.0032.because it represents the appropriate net work chosen in our work as indicated in Figure 1

The overriding thing to mention is that results that are related to error rates show that neural net works with non-ordered data proved to be efficient compared to neural net works with ordered data, because they provide an error rate of 0.0032 which is a very satisfying rate according to specialists. Hidden layer with five nodes also proved to be significant, because error rate changes in neural net work number 5 with non -ordered data compared to neural net work number 4 of the same rank.

Conclusion

To wrap it up, this research was an attempt to recognize Arabic minimal syllables automatically. It hinges upon three technological tools; Gold Wave, SFS and Neural Net Works, for processing because it is a technical work. It dealt with speech recognition which was used as a platform for the research, then talked about intelligence, types of intelligence and, neural networks, and then moved to syllables because they are the core of the research., syllable in Arabic, the structure of syllable in Arabic and syllable's recognition using neural net works were presented afterwards since one type of Arabic syllables was processed. The goal of this paper was to treat Arabic minimal syllable, so as to use Arabic in the field of artificial intelligence. It was through a modest experimentation, where the three aforementioned technological tools were used hinging upon the stratified random sample represented in 120 samples after being recorded by 40 Algerian speakers who were first divided into four groups, then were chosen randomly with different age and gender. The optimal neural net work is that of non-ordered data base, it consists of one layer, five nodes and 150 steps, because it has given a satisfying error rate and worked better than that with ordered data, in addition the principle of automatic recognition of one syllable is the same for other syllables and same even for other languages though different in their phonological system. To conclude we may say that Arabic phonological system is flexible and can be used easily in the field of artificial intelligence because of its limited number of vowels which can help in doing many combinations with consonants. And as a suggestion, firstly we recommend the use of SFS in sound processing because it is better than MATLAB in the sense that MATLAB requires a special training from the part of researcher while SFS does not

and secondly, the application of NNW with one hidden layer of five nodes and with non ordered data, because it can provide a good result in speech recognition. Finally it can be said that despite the great effort devoted to this humble research, it is considered as a first step to recognize one type of Arabic syllables automatically called minimal syllable, however much work still needs to be done for automatic recognition of all the types of syllables to expand the knowledge base, so that computer specialists will be able to join them and convert Arabic speech into manuscript.

About the Author:

Mohammed Dib is a Doctor (Maitre de conference B) at the University of Mascara in the English department, he studied at Tlemcen University, where he got his licence in English,his Magister and his doctorate in applied linguistics.. His area of interest is phonetics, phonology, signal processing, speech recognition, contrastive linguistics and automatic translation.

References

Abdelouhab.H. (1984) Introduction à La phonétique Orthophonique Arabe. collection Almoujtamaa Préface office des publications universitaires 1place centrale de Benaknoun. Alger

Anis.I. (1971)Alaswat Allughawia...fourth edition.published in AngloEgyptian library 165 avenue Mohammed Farid. Cairo.

Aljahid. Abu Otman Amru ibn Bahr (2002) died in Basra in 255 H Albayan wa Tabyin..Proofread by Darwish Jawidi. (1st ed.) Hilal house Beirut

Connor.J.D.O(1973)Phonetics.A simple and practical introduction to the nature and use of sound in language. .(1st ed) Cambridge university. F lesh.H. (2007)Nahwa Binaa Jadid.. edited translated and proofread by doctor Abd Sabour Shahin.first published in Jordan.

Gutman.R.D (2009)Chess Metaphors.Artificial Intelligence and the Human Mind- (Trans by Deborah Klosky).The MIT Press Cambridge, Massachusetts .London, England

Hagége. C. (1997) La Structure des Langues. 6eme édition collection fondée par Angoulvent encyclopédique.

Brown.D.H, (2007) Principles of language Learning and Language Teaching(.5 ed.) ISBN 0-13-199128-0

The Hutchinson Encyclopedia (1999) ed. published in by Helcon .ISBN 9781859862544

Joe.T. (May 1995) Speech Recognition Using Neural Net Works..CMU-CS-95-142. School of computer science Carnegie Mellon University.Pitsburg,Pennsylvania15213-3890

Kaddour,A.M(1999) Kaddour Mabadie allissaniat (2nd ed.) dar alfikr syria Beirut,Lebanon.

Lawrence. Rabiner Bing Hwang Juang Prentice (1993) Fundamentals of Speech Recognition... Hall International-Inc Published by PTR Prentice Hall,Inc .A.Simon and Schuster Company Englewood Cliffs,NewJersey 07632.

Mubarak,H..(1997)Fi asiwata azamania alwaqf fi lisaniat alklassikia. . Published in Rabat.

Peinado,A.M. Speech Recognition Over Digital Channels Robustness and Standards. University of Granada, 3Spain José C. Segura .University of Granada, Spain.

Roach.P.(1991).Introducing phonetics.. published by the penguin group.Ph Matthews Oxford concise dictionary of linguistics oxford university press.

Reinier. P. (2002)The Intelligent Ear on the Nature of Sound Perception. Professor Emeritus.Free University Amsterdam Lawrence Erlbaum Associates, Publishers Mahwah, New Jersey London.

Rumani. (1976) "Rissalat Nukat fi Ijaz Coran" (Trans as :A treaties on Expression Accuracy in the Miracle of Coran) proof read by Mohammed Khalaf Allah Ahmed and Mohamed Zaghloul Salam.(3rd ed.) Maarif house Egypt.

San,D. (2008) Syllable Structure. The Limits of Variation. Great Clarendon Street, Oxford ox2 6DPOxford University(1st ed.) published in the United States by Oxford university

W. Street Oxford Advanced learner's Encyclopedia dictionary(2010) .The ideal study dictionary for language and culture. Oxford university press Oxford .ox26Dp.new 8th edition.

WuChou,A. (2009) Minimum classification error(MCE) Approach in pattern recognition . labs research Avaya Inc USA.