Arabic Automatic Speech Recognition The Case of Minimal Syllable

DIB Mohammed M. A.
Université de Mascara

Abstract
This paper aims at shedding light upon a very complex problem which concerns speech recognition with regard to Arabic language trying to account for some techniques of implementing Arabic in the field of artificial intelligence
This work hinges first and foremost on three technological tools; Gold wave ,Mat lab and Neural Net Works to treat automatically the minimal syllable located in first, mid, and final position of three Arabic words recorded by thirty speakers of different age and sex.
Many experiments have been done in this work regarding the types of neural networks .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

1. Introduction

It is becoming something ridiculous 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 communication between man and machine.
My paper is under the title of "Automatic Treatment of Arabic language (Case of Minimal Syllable)" and there is no doubt that the spurs behind this work are: The sharp

deficiencies of the Arabic language as far as computer sciences are concerned, more simply put the problems that Arabic language is facing nowadays with regard to globalization which tries to put it back in the black despite its potential words, the number of its users and its system in terms of flexibility and exactness.

My present paper aims at shedding light upon a very complex problem which concerns speech recognition and remains ambiguous with regard to Arabic language. In this account I try to identify limitations that point the way to this research relying upon Neural Net Works as a 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, automatic 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.

2. 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[1]. In other words an analog signal is converted into an acoustic signal or digital one to get classified decisions, or

[1] Wu Chou Avaya. Minimum classification error(MCE) Approach in pattern recognition.p:14

a variable stimulus is transformed into a constant response.[2]

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.[3]

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[4]. Automatic speech recognition requires knowledge of many fields such as signal processing, acoustic phonetics, patterns recognition, communication, theory of information, physiology

Intelligence

No exact definition is found for the word intelligence, however some attempts were made to make the term more clear. It is the ability to understand and to think[5]. It is the ability learn, comprehend and to think (Longman advanced dictionary date)[6]. 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

[2] .R. Plomp. The Intelligent Ear on the Nature of Sound Perception p:94

[3] L. Rabiner &B. H. Juang Fundamentals of Speech Recognition.p:4

[4] Antonio M. Peinado. Speech Recognition Over Digital Channels Robustness and Standards. P:1

[5] Oxford Advanced learner's Encyclopedia dictionary.

[6] Longman advanced dictionary

intelligence, kinesthetic bodily intelligence, personal intelligence and interpersonal intelligence[7].

As for artificial intelligence, it can be said that it is the reproduction of all types of intelligence using artificial means such as Neural Net Works. 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, linguistic field and understood the process of language and therefore dealt with NLP. [8]Moreover artificial intelligence tries to realize tasks similar of that used by human intelligence. It entails robot behavior, language comprehension, patterns recognition, and knowledge representation.[9]

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 NNW.

NNW 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

[7] .H. Douglas .Principles of language Learning and Language Teaching p. 100

[8]Wafaa Niar .Dinedane. L'Intelligence Artificielle Elément de Base. p :13

[9] The Hutchinson Encyclopedia.

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.[10] They represent many mathematical models of human brain's functions such as: Comprehension, calculation and memorization.

There are many different types of NNW, however they have same four basic attributes which can be summarized as a set of processing units, a set of communication, a computing procedure, and a training procedure. [11]

Syllable

No definition is completely satisfactory for 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 are not considered as syllables are to be avoided[12] .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:IN**]and the pulse theory cannot decide whether the word *beer* [**biÄ**] contains two syllables in American pronunciation. This question ship generates some doubts on whether syllables are linguistic units or not. Chomsky, Halle, Steriade, Gimson, Belvins do not consider them as phonological units.[13] Despite all this syllable appears clearly in some cases for example people agree that the word Canada

[10]Op cite

[11] Joe. Tebelskis. Speech Recognition Using Neural Net Works. p:28

[12] San Duanmu. Syllable Structure. p :36

[13] Ibid

entails three syllables in contrast the word America comprises four syllables.[14]

The Syllable in Arabic

The syllable in Arabic always starts with a consonant and ends either with a vowel and is called the open syllable or with a consonant and is called the close syllable. This means that the word in Arabic never starts with a consonant cluster, more simply put Arabic rejects the starting of three connected plosive consonants. 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.[15]

Generally speaking the syllable is an association between a consonant and a vowel. Roman Jacobson 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.[16] Ampere Crambe, however, argues that speech hinges upon breathing, and air exhaling is similar to pulsation, each muscle contraction with the increase resulting from the air pressure forms a chest pulse, and

[14] Op cite
[15] Henry Flesh. Nahwa binaa jadid.p:43
[16] H. Abdelwahab. introduction à la phonétique orthophonique Arabe.p:27

each chest pulse in its turn forms a syllable. The pattern of chest pulses is the basis of human language.[17] Some sources said that the dividing speech into syllables goes back to a long time with regard to Arabs and goes back to the period where Arabic language was an oral language and relied only on the listening process to transfer literature and arts. Aljahid one of the Arab grammarians used the term syllabification which means segmentation of speech. He said that the sound is a device of speech whose role is syllabification and connection.[18]

Connected speech encompasses syllables that bear the phonotactic of a particular language. It has been mentioned earlier that syllables may be divided into open and closed ones. Open syllables are those which end with short or long vowel, but closed syllable are those which end with a consonant with an absence of mark referred to as sukun. The tri lateral root verb (fataha) contains three open syllables, yet the noun (fathun) comprises two closed syllables: /fat/ and/ hun/.[19]

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 and relies upon the respiratory rhythm. The minimal syllable in Arabic is formed from one consonant followed by a short or a long

[17] M. Hanun. Fi asiwata azamania alwaqf fi lisaniat alklassikiap:65
[18] Aljahid. albayan wa Tabyin p:79
[19]. I. Annis. Alaswat allughawia p:162

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.[20] 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/. Syllabic consonant appear because of the deletion of the weak vowel schwa.[21]

Types of syllables

There is a controversy about the number of syllables in Arabic, Some linguists argue that there are six types, but others said there are only five. They are as follows:

1. The minimal syllable: It consists of a consonant and a vowel. They are meaningful linguistic units, they consist in prepositions e.g. bi ,fi,li…etc.
2. Closed long syllable: It consists of a consonant, a vowel and a consonant.eg. mithl (like) min (from) bal(rather)

[20] C. Hagége. La structure des langues p:24

[21] Peter Roach. Introducing phonetics.p:106

3. Open long syllable: It consists of a consonant and two vowels .e.g. maa, haa Two forms related to the pause.
4. Long syllable closed with a consonant: A consonant+two vowels +a consonant e.g. kaan (was)
5. Long syllable closed with two consonants*:* A consonant + a vowel + two consonants e.g. karb, fadl[22]

Data base

Arabic language contains 28 consonants, 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, and the absence of the mark referred to as sukun.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 syllable. Short vowels, long vowels, and minimal syllables are displayed in the following tables.

[22] A. M. Kaddour. Mabadie allissaniat .p:75

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:**

Table of vowels

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

Table of minimal syllables

As far as knowledge base is concerned, it should be noted that three words containing the minimal syllable/ **ʔa/** located in first, mid and final position were recorded by 30 speakers then segmented and as a result 90 syllables were obtained.

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

Syllables 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, 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 analog signal is converted into 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. And then to 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. formants F1 F2 F3 F4, periodicity, and amplitude, then try to train the system to make an association between input and output.

Before all, it should be noted the nature of this work requires the use of 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 garcon (boy). As far as the letter *f* is concerned, there are two: one means final and the other means femme (woman). As for the numbers, they represent speakers. The words mentioned above were recorded by 30 informants of different sex and age; 5 men, 5 women, 5 boys and 5 girls. Then segment to

have the minimal syllable referred to as ***hamza*** located in first, mid and final positions.

Two charts are used for the process of recognition; one represents the referential sample and the other the test sample. We mean by referential sample input and output of sounds. 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.

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	721	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**
Aff5	**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 showing the characteristics of minimal syllable located in first, medial and final position
Referential samples

Test sample

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	0.1919	0.0001667

	NNW with ordered data			
Number Of NNW	**Number of layers**	**Number of nodes**	**Number of steps**	**Error rate**
1	**1**	**2**	**150**	**0.9783**
2	**1**	**3**	**150**	**0.9703**
3	**1**	**4**	**150**	**0.9791**
3	**1**	**5**	**150**	**0.9791**
	NNW with non- ordered data			
Number Of NNW	**Number of layers**	**Number of nodes**	**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 showing types of Neural Net Works with ordered and non- ordered data

Many experiments represented in eight Neural Net Works have been done in this work to show which one can bring better results; four Neural Net Works with ordered data, while four other Neural Net Works with non ordered data. As far as NNW with ordered data are concerned, we have the following:

Neural Net Work number 1 consists of one layer, two nodes , 150 steps and the error rate is 0.9783.the second net is of 1layer, three nodes,150 steps and error rate of 0.9703.the third net is of one layer, four nodes ,150 steps and error rate 0.9791.Net work number four ,however, contains 1layer five nodes ,150 steps and error rate of 0.9791.

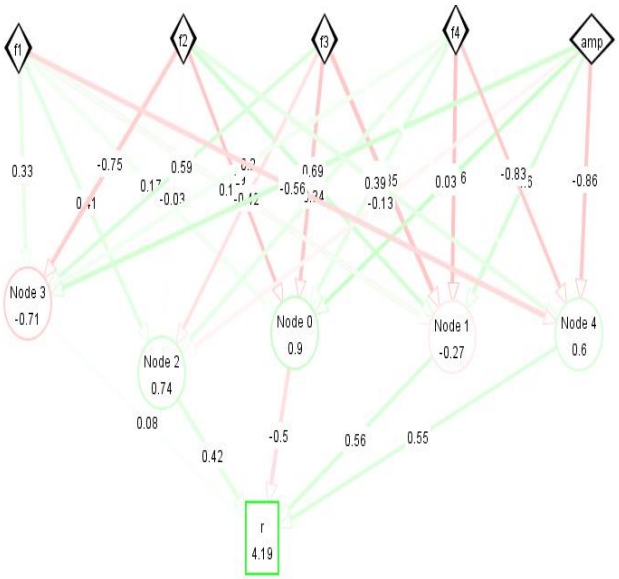
As regard NNW with non ordered data they are displayed as follows:

NNW 5 contains 1layer, two nodes, 150 steps and error rate 0.0074.
NNW6 has 1layer, three nodes, 150 steps and error rate0.0060.
NNW7 is of 1 layer, four nodes, 150 steps and error rate of 0.0047, while NNW 8 has an error rate of 0.0032 with one layer five nodes and 150 steps.
The optimal NNW is Net work number 8 with one layer five nodes and 150 steps and error rate of 0.0032.

Conclusion

This research is an attempt to recognize Arabic minimal syllables automatically relying upon Neural Net Works; one tool of scientific matters which was applied hinging on the random simple sample represented in 90 samples recorded by 30speakers from different age and sex. The optimal Neural Net Work is that of non-ordered data base, it consists of one layer, five nodes and 150 steps, because it gives a satisfying error rate and works better than that with ordered data, more simply put the principle of automatic recognition of one syllable is the same for the other syllables and same even for other languages though different in their phonological system.
It can be safely 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.
The following design represents the appropriate net work chosen in our work. It comprises one input layer of sound

characteristics represented in F1, F2, F3, F4 and amplitude, one hidden layer of five nodes and one output layer called result.



150 Steps .Error rate 0.0032

References

Ahmed Mohammed Kaddour (1999) Mabadie allissaniat second edition, Beirut,Lebanon.

Aljahid Abu Amru ibn Bahr died in Basra in 255 H albayan wa Tabyin..proofread by Darwish Jawidi.

Antonio M. Peinado (2002) Speech Recognition Over Digital Channels Robustness and Standards. University of Granada, Spain José C. Segura .University of Granada, Spain.

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

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

Henry Flesh (2007) Nahwa binaa jadid.. edited translated and proofread by doctor Abd Sabour Shahin.first published in Jordan.

Ibrahim Aniss(1971) Alaswat allughawia...fourth edition.published in AngloEgyptian library 165 avenue Mohammed Farid. Cairo.

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

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,NewJersey07632.

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

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

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

San. Duanmu (2008)Syllable Structure. The Limits of Variation. Great Clarendon Street, Oxford ox2 6DPOxford University Press is a department of the

University of Oxford.It furthers the University's objective of excellence in research, scholar ship,and education by publishing worldwide in Oxford New York
The Hutchinson Encyclopedia 1999 edition
W.N. Dinedane(1995) L'Intelligence Artificielle Elément de Base. . office des publications Universitaires codification 2.10.3997
W. Street (2010) Oxford Advanced learner's Encyclopedia dictionary .The ideal study dictionary for language and culture. Oxford University Press Oxford .ox 26 Dp. new 8th edition.
Wu Chou Avaya.(2009) Minimum classification error(MCE) Approach in pattern recognition. labs research Avaya Inc USA.