Synthesis and Characterization of Some **Age classification using speech signal**

Khalil Al-saif[1] , Shymaa Alhyali[2]

[1]*Computer science dept , Mosul Univ. , Mosul , Iraq*
[2] *Computer center , Mosul Univ. , Mosul , Iraq*

(Received 3 / 11 / 2008 , Accepted 25 / 10 /2009)

Abstract

In this research the relation between the age and speech signal was studied depending on three ways :
1-Time domain .
2-Frequency domain .
3-Polynomial technique .
The study shows that the speech signal after applying the primacy process on it, using a packet of different size, then extracting the common feature for different age and different sex to be adopted at the recognition stages .
The study show the classes of ages, when the time domain was use, give limited rang of time can be used as an introduction to age suggestion of the speaker when recognizing the speech signal .
As well as a conclusion got that the retrieval signal (when the polynomial technique used) does not effected when the polynomial degree goes over 15 .

1. Introduction :

Sound is a result of changes of pressure in the air outside our ears[1]. Acoustic instruments generally produce sounds when some part of the instruments is either stuck, plucked, bowed or blown into [2][3].
Sound is what we experience when the ear reacts to a certain range of vibrations. These vibrations themselves can also be sound. the average person can hear sound from about 20Hz to about 20,000Hz. the upper frequency limit will drop with age .[2] and may reach (20Hz-17KHz) [1][4].
Of course, pressure variations with frequencies lower than 20 Hz should also be regarded as sound. They are normally referred to as infrasound.
Sounds with frequencies higher than 20 kHz lie above the audible region and are referred to as ultrasound[4], Fig (1) show the movement of vibrication in sound wave .

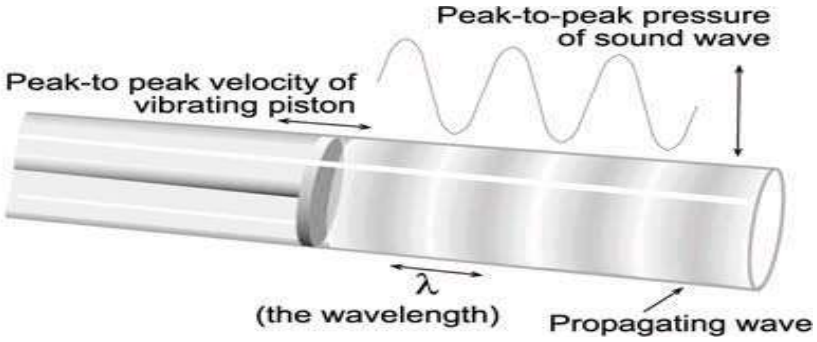


Fig (1) The movement of vibrication in sound wave

Technically, sound consists of pressure waves that move through a compressible medium. The precise mechanics vary somewhat according to the material, but the broad principles are the same. Molecules, whether of air, water, or metal, like to keep on average equally for apart from all of their neighbors. So whenever the molecules in one area are turn compresses their neighbors [5], as shown in Fig (2).

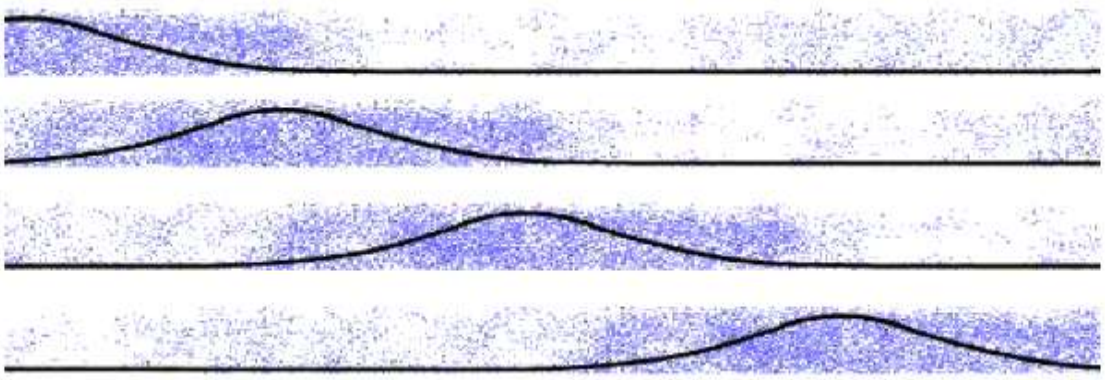


Fig (2) A moving pressure wave

The dark areas in Fig (2) indicate where molecules are squeezed close together (these are areas of high pressure). The lighter areas indicate where molecules are relatively sparse. When the tightly compressed molecules in the first image force themselves apart, they end up compressing their

neighbors, who in turn will force apart to compress their neighbors, and so on [5].

In order to understand the sound we must understand something about the signals [3]:

Sound Frequency:

The vibration rate of a sound is called its frequency the higher the frequency is the higher the pitch. Frequency is often measured in units called Hertz (Hz) [2].

There is a physical element attached with the frequency is (pitch) [2], pitch correspond to the frequency, but there are two specific limitations on the human ear [1]:-

1-First is that the range of sounds that we can hear is limited.

2-Second limitation is that with complex waves of more than one frequency, the frequencies are not normally perceived independently, but as being one particular pitch with timbre which is determined by the other frequencies present.

Sound Amplitude:

The continuous rise and fall in pressure creates a *wave* of sound. The amount of change in air pressure, with respect to normal atmospheric pressure, is called the wave's amplitude. We most commonly use the term "amplitude" to refer to the peak amplitude, the greatest change in pressure is achieved by the wave [6][3].

Loudness corresponds roughly to the amplitude of the note. It is not quite the same because the sensitivity of the human ear is different at different frequencies, and hence the perceived loudness of a sound depends on both the frequency and amplitude [1], Fig (3) show the difference between amplitude and frequency .

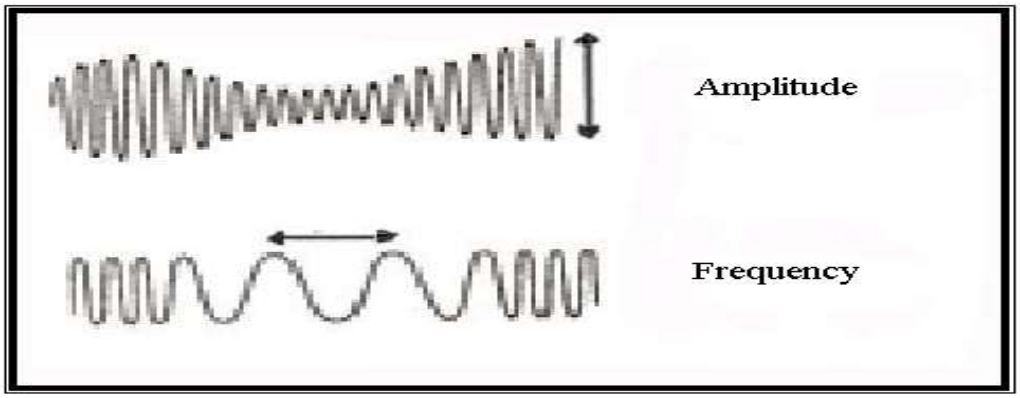


Fig (3) The difference between the amplitude and frequency

Sound Timbre:

Is a French word that means "tone color". It is pronounced: tam' ber. Timbre is the quality of sound which allows us to distinguish between different sound sources producing sound at the same pitch [2].

The vibration of sound waves is quite complex; most sounds vibrate at several frequencies simultaneously. The additional frequencies are called overtones or harmonics. The relative strength of these overtones helps determine a sound's timbre [2].

There is no simple physical property which corresponds to this component. It is a product of many different things [1].

Sound Envelope:

Another important factor in the nearly infinite variety of sounds is the change in over-all amplitude of a sound over the course of its duration [6], The four portions of the envelope are [1]:-

Attack:

Is the period from the onset of the note to its peak volume.

Decay:

Represents a period when the peak volume decreases to another level which may be held for a period of time.

Sustain:

The period when the second volume is held is the sustain.

Release:

The final decrease in volume from the sustain level to nothing is the release.Fig (4) show the four portions in sound wave.

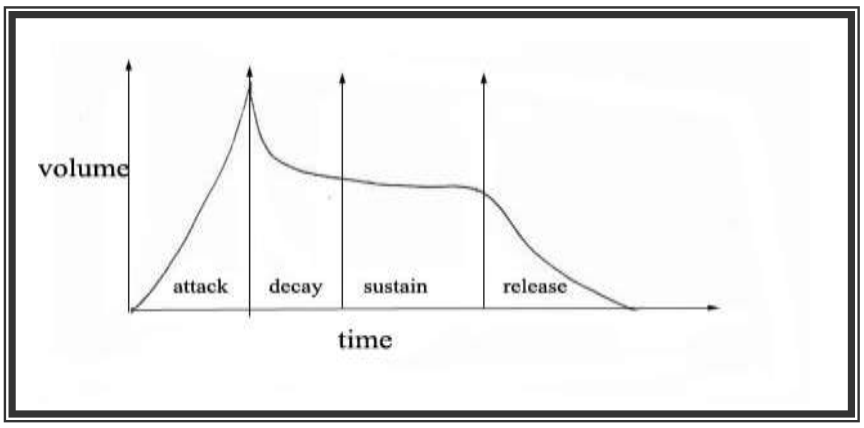


Fig (4) Envelope of the sound wave

2. Digital sound :

When recording audio comes in and out of a computer, process of converting audio data from analog to digital and digital to analog needed . This is an extremely weak link in the recording process and can have a drastic effect on the sound quality. Do you ever wonder why digital sound is sometimes "sharp" or "harsh"? Repeated analog to digital and digital to analog conversion is what causes this unsatisfying sound. The remedy is a superb A to D converter.

Flexibility
Convenience
Clarity
Longevity
Affordability

3. Digital Sound Recording:

When the computer receives a sound from a microphone connected to sound card, the microphone transforms the vibration of air pressure into a vibration in voltage in a form of an analog symmetric signal works by sound power at the receiving of this signal by measuring and transferring it into parts called sample and then into a series of digits in a process called quantization. These digits are then transferred into binary form. These samples are stored inside the computer on the hard disk as a binary digital form [7][8].

The Analog-to-Digital Converter (ADC) is the device in which both quantization and binary coding of the sampled signal take place.This electronic circuit will help the operation of recording audio and this circuit will appear on the sound card [9] as shown in Fig (5) .

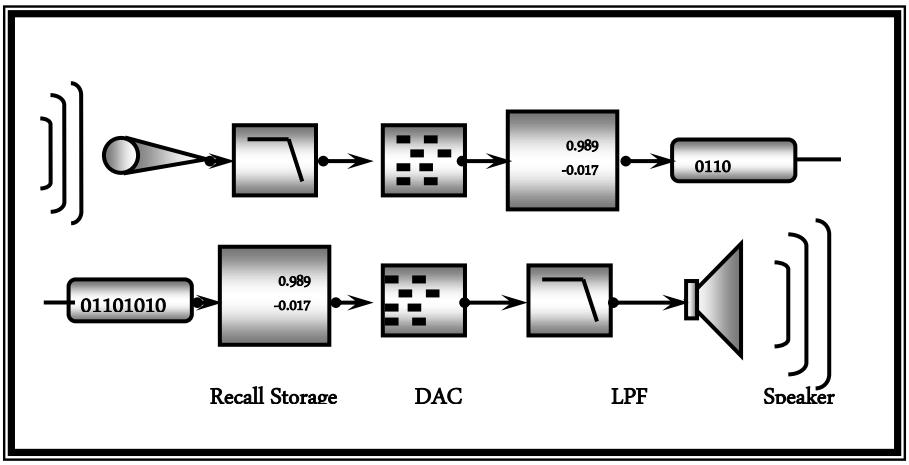


Fig (5) Convert the sound from the analog format to digital format and vice versa

In the operation of recording there are many important elements which must record with it [10]:

3.1 Number of Channels:

stereo files work with two separate audio channels, and they are usually about twice the size of mono files. It is theoretically possible to record any number of channels, with a concomitant increase in file size. For example, a file intended to be played on a surround-sound system may record seven or more channels, and will thus be about seven times the size of a mono version of the same sound. [11][10].

3.2 Sampling Rate:

Most digital sound files save information as a long series of sound samples, in the same way a film saves moving pictures as a series of still images. The quality of a sound file can be increased by taking more of these samples in the same amount of time; that is, increasing the "sampling rate". This has the effect, however, of increasing the file size[10], Fig (6) show the sampling of Sound wave.

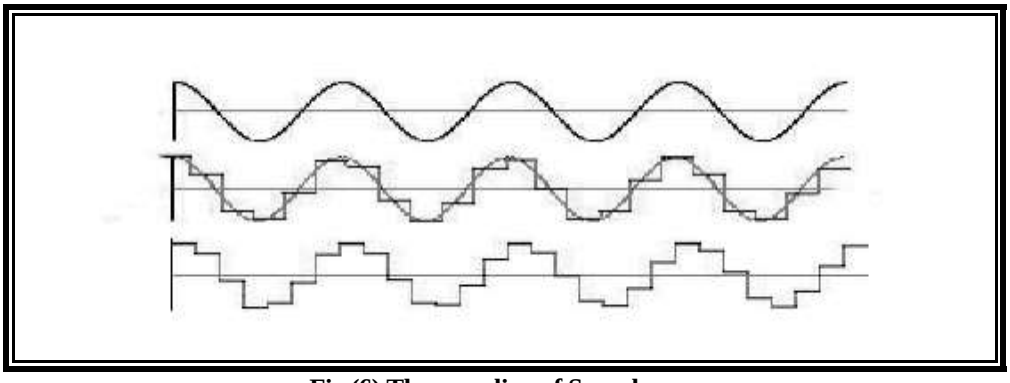


Fig (6) The sampling of Sound wave

Converting an analog signal into digital signal causes the loss of part of the information of the signal, to avoid this state and to get high precision in the process of sampling , rate of sample per second is taken to be at least double the highest frequency in the analog signal [2][12][13], that is:

$Fs \geq 2f$

Where :

Fs=sampling rate.

f=the highest frequency.

The Nyquist-Shanon sampling theorem, a fundamental theorem of signal processing, states that a sampled signal

cannot unambiguously represent signal components with frequencies above half the sampling frequency[12][14].
We use a rate of 44,100 Hz measurements (samples) of sound per second in our professional equipment; this sampling rate of 44,100 Hz (or 44.1 kHz) is called CD quality recording because it is used in commercial CDs. Other common sampling rates are 22,050 Hz (for lower quality, multimedia files) and 48,000 Hz for DAT recorders [2][3].

3.3 Bits per Sample:

Commonly, either 8 or 16 bits are used to represent each sample. Using 16 bits provide for much better quality, but produces files twice as large as 8-bit files [10].
Therefore, sound quantized with 8-bit numerical precision will have a best case (SQNR) Signal-to-Quantization Noise Ratio of about 48 dB. This is adequate for cases where fidelity is not important, but is certainly not desirable for music or other critical purposes. Sound sampled with 16-bit precision ("CD-quality") has a SQNR of 96 dB, which is quite good--much better than traditional tape recording [6].

4. The aim of feature extraction :

The essential application of the algorithm feature extraction is to represent the speech signal and to reduce the ratio of data entary when the processing stage start, so the less number of features may prevent the data repeated, and hold on the initial information needed in next processing stage. So the analyzing operation process may be closer to the exact age of the speakers [15].

5. Extract the speech feature :

The operation of extract feature to represent the speech signal, can be done by converting the speech wave to their parameters, to reduce the ratio of the data required, so the parameters will aid to reconstruct that speech .
Due to that, the classification of the age is depend in the feature parameters .
Therefore it is difficult to predict the speech signal before classified it to their frame numbers and features, because of the stability of the speech signal within the short duration of time. The process is reveal numerous feature vector which is equivalent to the number of the signal frame. Fig (7) explain briefly the subsystem of the feature extraction for human speech.

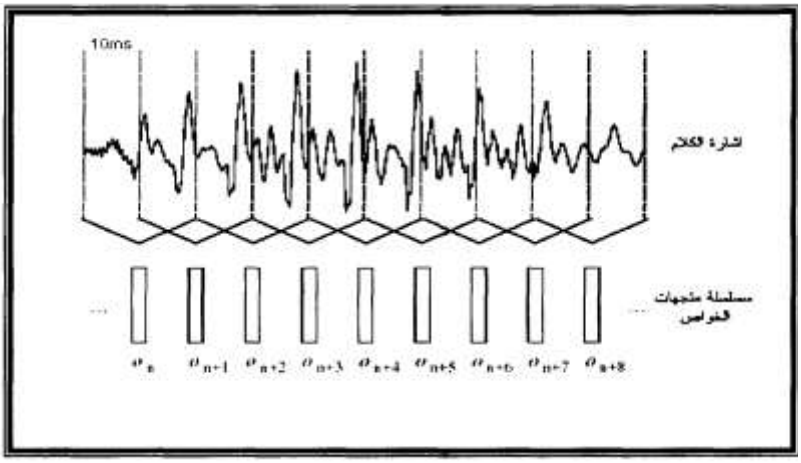


Fig (7) subsystem for feature extraction of human speech .

In spite of that there is no agreement on ideal feature for speech signal, but it's better that the extract feature of the speech recognizes as :
1. It must permit for the automatic system to distinguish between the similar sounds for speech signal .
2. Also it permit for automatic construction for sound model, (i.e.without needing for whole signal parameters), from it's limit of training data .
3.Given a different statistical data must be provided with different speakers on different environments .
In order to extract the speech features, (dealing with speech signals), speech must be adopted using time domain or frequency domain .

6. Fitting use Polynomial techniques :

The least square method is one of the best way to get a root locus of a straight line from a group of events (x,y) . It works to minimize the square error (i.e. to its minimal value) , the procedure is explained in equ (1)[15]:

$$y=\sum_{i=1}^{m}\left(y_i-a_0-a_1x_i-a_2x_i^2-.....-a_nx_i^n\right)^2 \quad \text{.....(1)}$$

m: no of events .
A:the value of elements that specify the equation.
N:the degree of equation .
Taking the partial derivative of equation (1) with respect to all the parameters, then equalize them to zero as shown in equations (2,3,4):

$$\frac{\partial s}{\partial a_0}=-2\sum\left(y_i-a_0-a_1x_i-a_2x_i^2-.....-a_nx_i^n\right)=0. \quad \text{....(2)}$$

$$\frac{\partial s}{\partial a_1}=-2\sum x_i\left(y_i-a_0-a_1x_i-a_2x_i^2-.....-a_nx_i^n\right)=0. \quad \text{....(3)}$$

$$\frac{\partial s}{\partial a_n}=-2\sum x_i^n\left(y_i-a_0-a_1x_i-a_2x_i^2-.....-a_nx_i^n\right)=0. \quad \text{.(4)}$$

rearrange the polynomial to find :
Then a number of equation will be obtained equal to the number of variables in the equation .
By Solving the set of equations (5,6,7) the variables (a0,a1,a2…) will be known .

$$a_0 m + a_1 \sum x_i + a_2 \sum x_i^2 + \ldots\ldots + a_n \sum x_i^n = \sum y_i \quad . \ (5)$$

$$a_0 \sum x_i + a_1 \sum x_i^2 + \ldots\ldots + a_n \sum x_i^{n+1} = \sum x_i y_i \quad \ldots\ldots (6)$$

$$a_0 \sum x_i^n + a_1 \sum x_i^{n+1} + \ldots\ldots + a_n \sum x_i^{2n} = \sum x^n{}_i y_i \quad \ldots\ldots (7)$$

7. Speech signal representation using :

7.1Time domain :

In this type of domain, the feature extraction of speech signal directs and don't need as a sample to represent the signal, so the process of a signal in time domain easily and faster than the other domains .

7.2 Frequency domain :

This type of domain is widely used in feature extraction, when the sound signals is treated, it seen it contains a useful constants coefficient which will be used for distinguish the speech signals, and they can be extracted from spectral analyses .

The process of conversion from domain to another is done by using fast fourier transform and inverse fast fourier transform (fft and ifft), also it is known as a conversion between time domain and frequency domain (vise versa). This transform can be introduced by a pair of equations(8),(9):

$$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x) \exp[-j2\pi ux / N] \quad \ldots. (8)$$

$$f(x) = \sum_{u=0}^{N-1} F(u) \exp[j2\pi ux / N] \quad \ldots (9)$$

Where :

f(x):the signal in the time domain .

F(u):the signal in the frequency domain .

N: number of sample in the signal .

J: the square root of (-1) .

8. Subjective tests:

The subjective tests use human listening in the valuation process, it simplifies the direct valuation test of the noise degree existing in the speech during the listening number of persons to the original signal and retriever signal, then to compare between them and registrate these notes (via some fixed points) . And one of the famous measurement which it works on this principle known as (Mos:Mean opinion score) which it represented on five levels, that from it, thus a limitation can be made for the received speech quality, Table (1) show the consideration opinion of the listener to input and output speech .

Table (1) Speech Quality Ranks

Rank	Speech quality	Distortion level
1	Excellent	The speech quality is excellent similar person speaking with another
2	Good	The speech which is similar to the speaking in the telephone
3	Middle	The speech is understand but is not as excellent quality
4	Acceptable	We can understand the speech but we can't know who is the speaker
5	Bad	We can understand the speech and also who is the speaker

9. Result discussion :

The tables and graphics which were given in attached appendix show, it is clearly to distinguish between the voice of different ages so the table gave the signal parameter in different domain and with different ages and sex.

Results :

When the sound was studied for different age of persons, the following points appear by applying the three approaches given in this research :

-The practical application using (1000) sampling rate to measure the retriever quality is the best result.

-The quality of retriever speech, by using the suggested algorithm is efficient, the test use six different appreciator and individual from each other.

-Using Fourier transform to apply the suggested algorithm doesn't lead to an excellent results, even it is faster to treat because it lead to deal with the Fourier transform parameters instead of the whole signal parameters .

-Fitting the signal using polynomial does not effected over the degree of(15).

In the appendix the coefficients of the speech signal was divide in three figures to give wide idea about them (the three approaches time , frequency and polynomial), while table (A1) of the appendix show different approaches to represent the speech signal .

Also the appendix give clear figures regarding speech signal (coefficient) of different genders in different ages , The figures were given in different approach (time / frequency / polynomial) covered in parallel the requirement needed to distinguish between the ages of different gender .

Appendix (A)

Table (A1) the speech coefficient (speech coding)

No	Some samples for recorded speech.	The values obtain when using the coefficient of curve fitting .	The samples that align the array of Fourier transform.	The real values of samples that represent the speech.
1	0.015076	0.015017	-0.47791	-0.47791
2	-0.0016479	-0.0010193	11.644 + 4.7273i	11.644
3	-0.011139	-0.013947	-1.2626 + 4.4283i	-1.2626
4	0.020355	0.026529	4.9326 - 2.4489i	4.9326
5	-0.021942	-0.026171	-0.15062 + 0.31081i	-0.15062
6	0.010712	-0.0013651	-0.32889 - 0.66732i	-0.32889
7	0.018585	0.056012	-0.13038 + 0.051623i	-0.13038
8	0.11774	0.067315	0.068599 - 0.057329i	0.068599
9	0.0065613	0.055132	1.2288 + 0.33994i	1.2288
10	0.12296	0.059137	0.57189 - 0.24444i	0.57189
11	-0.046204	0.062195	-0.11026 + 0.14016i	-0.11026
12	0.17053	0.034287	-0.0020457 + 0.017449i	-0.0020457
13	-0.10895	-0.0042208	-0.017158 + 0.17372i	-0.017158
14	0.031769	-0.0061431	-0.21225 - 0.48036i	-0.21225
15	0.042053	0.026664	1.0829 - 1.0307i	1.0829
16	-0.0085144	0.034941	0.31897 - 0.28652i	0.31897
17	0.051849	-0.0050792	0.055854 - 0.22975i	0.055854
18	-0.073517	-0.017547	-1.4714 - 1.0962i	-1.4714
19	0.10376	0.06405	1.6069 - 0.70553i	1.6069
20	0.10651	0.12581	2.4737 - 0.37773i	2.4737
21	0.027466	0.021374	-1.7259 - 0.70338i	-1.7259
22	0.028473	0.0296	0.269 - 5.9495i	0.269
23	-0.0065308	-0.0066245	32.636 - 7.8824i	32.636

Table (A2) coefficient of speech after retrieved

No	The values obtained when using the coefficient of curve fitting.	The values obtained after apply inverse Fourier transform .	The values obtained by taking the real part of inverse Fourier transform.
1	-0.47457	2.2174	2.2174
2	11.604	1.8711 - 0.17414i	1.8711
3	-1.0521	1.5321 - 0.39852i	1.5321
4	4.3051	1.1497 - 0.52845i	1.1497
5	0.96141	0.57343 - 0.69193i	0.57343
6	-1.3749	0.51353 - 1.0641i	0.51353
7	-0.030829	0.090294 - 1.4354i	0.090294
8	0.93949	-0.56509 - 0.94444i	-0.56509
9	0.54461	-1.1447 - 0.51486i	-1.1447
10	0.29379	-1.5676 - 0.2623i	-1.5676
11	0.42252	-1.8352 - 0.12269i	-1.8352
12	0.093504	-1.9637 - 0.035958i	-1.9637
13	-0.43828	-1.9637 + 0.035958i	-1.9637
14	-0.064117	-1.8352 + 0.12269i	-1.8352
15	0.96309	-1.5676 + 0.2623i	-1.5676
16	0.85991	-1.1447 + 0.51486i	-1.1447
17	-0.69169	-0.56509 + 0.94444i	-0.56509
18	-0.99435	0.090294 + 1.4354i	0.090294
19	1.4888	0.51353 + 1.0641i	0.51353
20	2.4409	0.57343 + 0.69193i	0.57343
21	-1.6934	1.1497 + 0.52845i	1.1497
22	0.2598	1.5321 + 0.39852i	1.5321
23	32.637	1.8711 + 0.17414i	1.8711

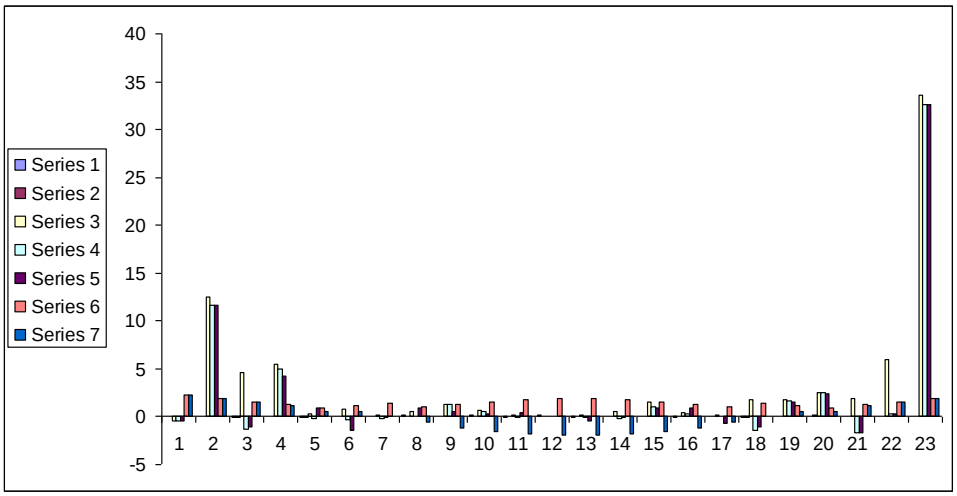
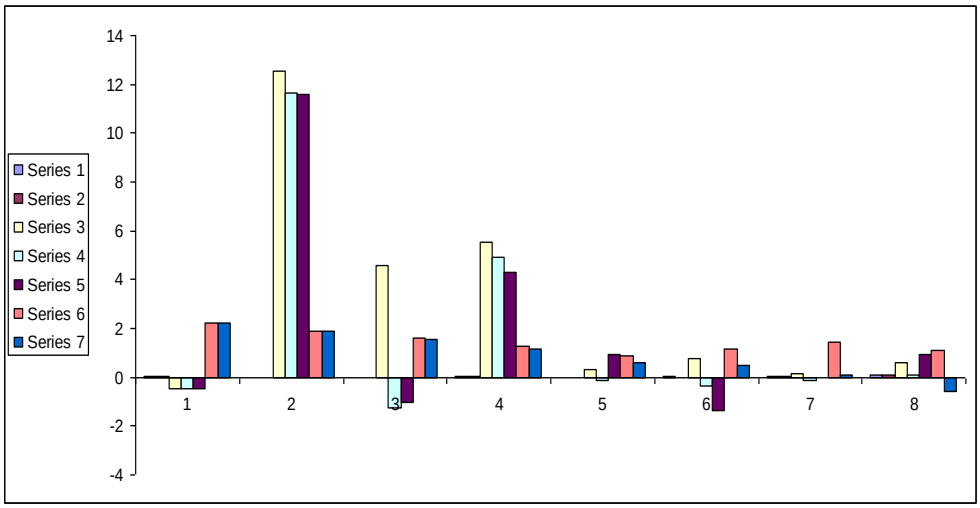
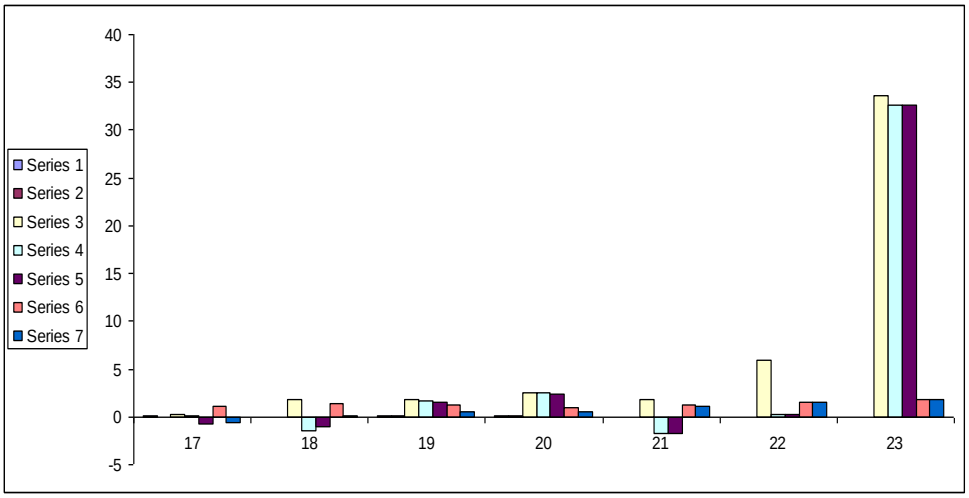


Fig (A1): the speech coefficient vs. the amplitude for (age=20 years and above) (sex=male)





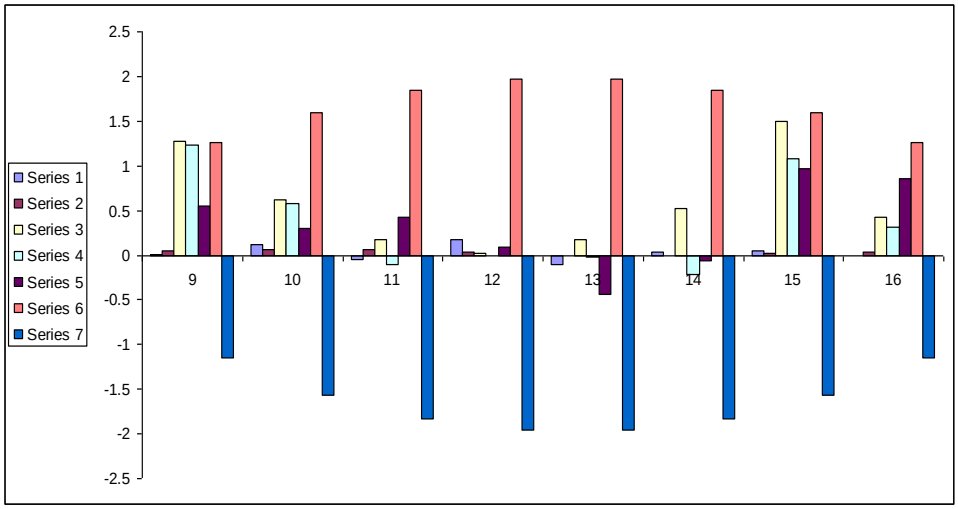


Fig (A1,b,c,d) coefficient of the speech (1-8) (9-16) (17-23)

Male sound

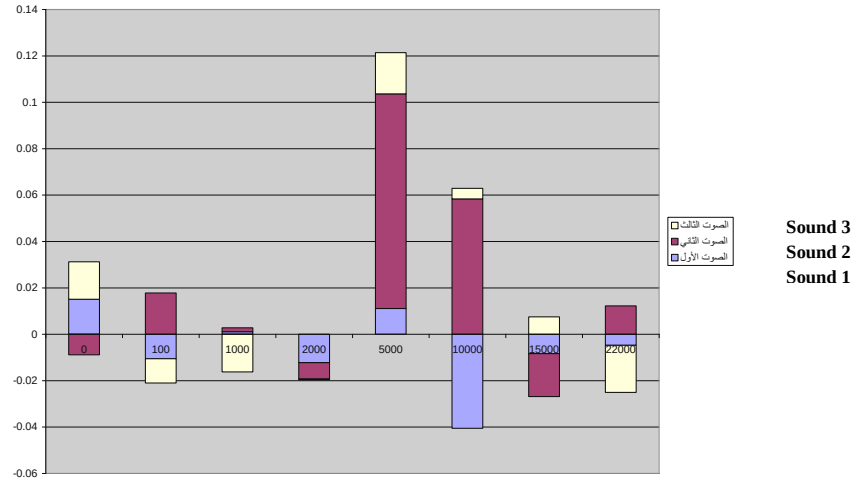


Fig (A2) the coefficient of mail sound in time domain for age (20 years and above).

Female sound

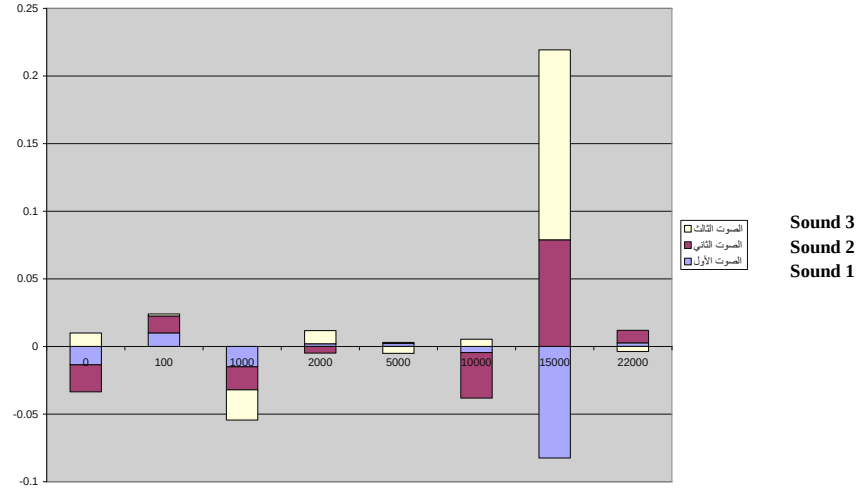


Fig (A3) the coefficient of female sound in time domain for age (20 years and above).

Male sound

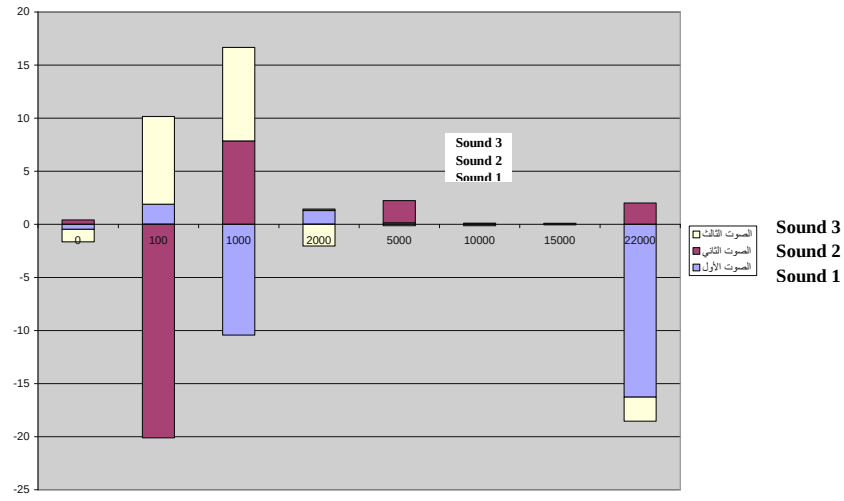


Fig (A4) the coefficient of mail sound in frequency domain for age (20 years and above).

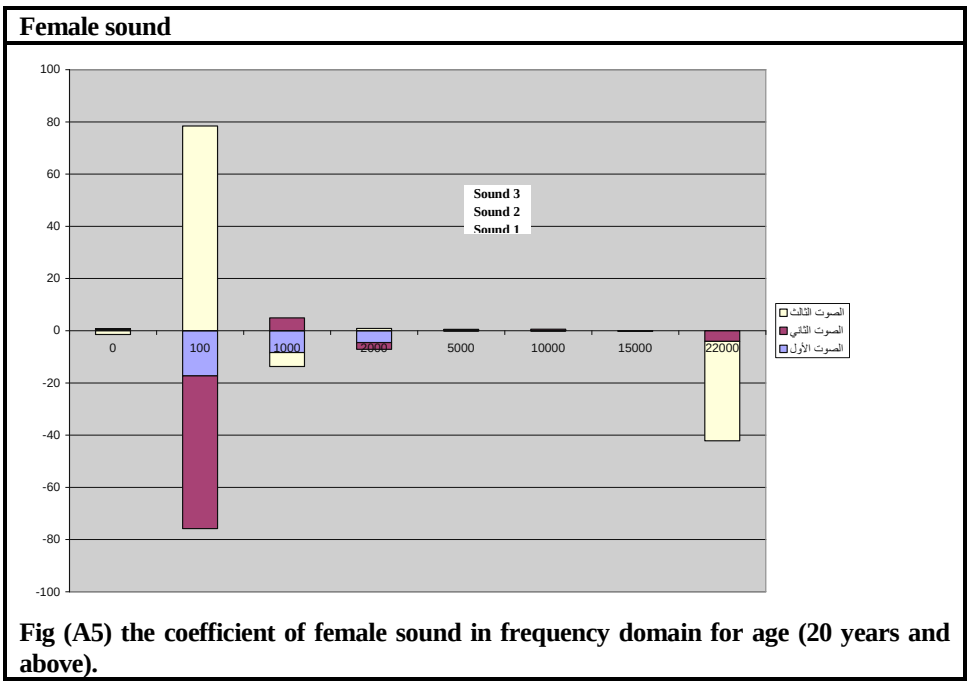


Fig (A5) the coefficient of female sound in frequency domain for age (20 years and above).

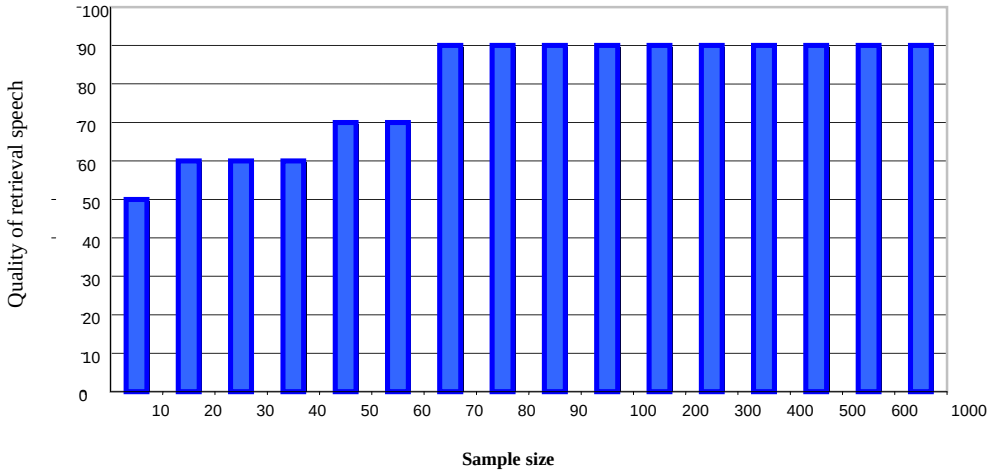


Fig (A6) the quality of the retrieval signal depending on time domain as a size of sample different .

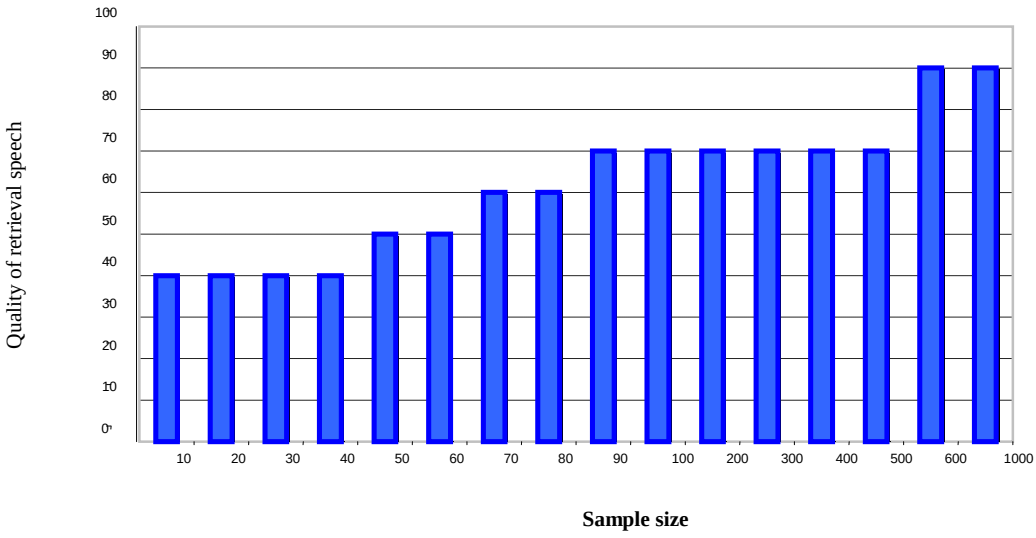


Fig (A7) the quality of the retrieval signal depending on frequency domain as a size of sample different.

Less than 10 years

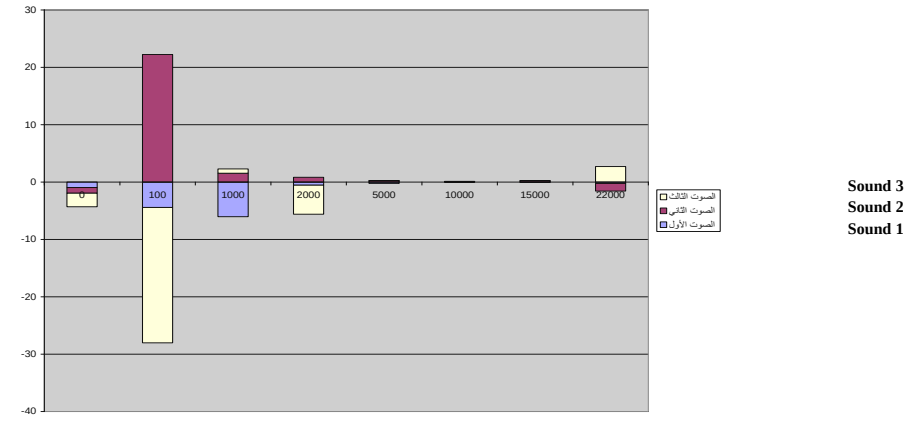


Fig (A8) the coefficient of male speech in frequency domain for age (less than 10 years).

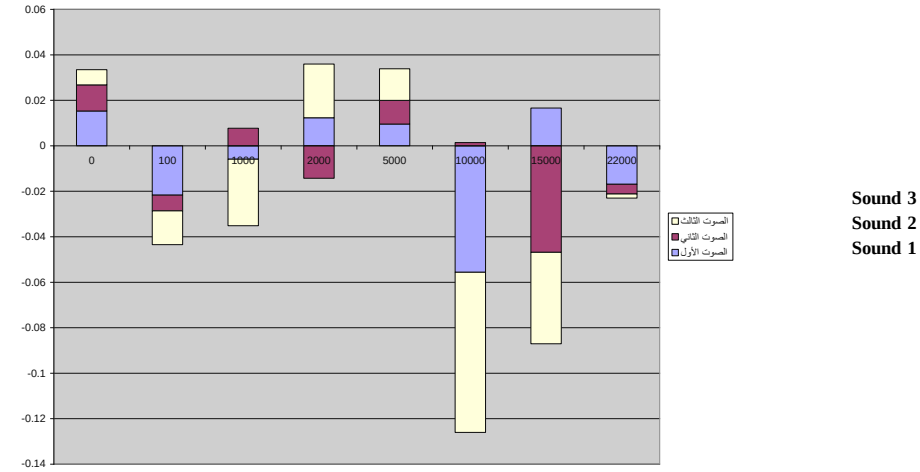


Fig (A9) the coefficient of male speech in time domain for age (less than 10 years).

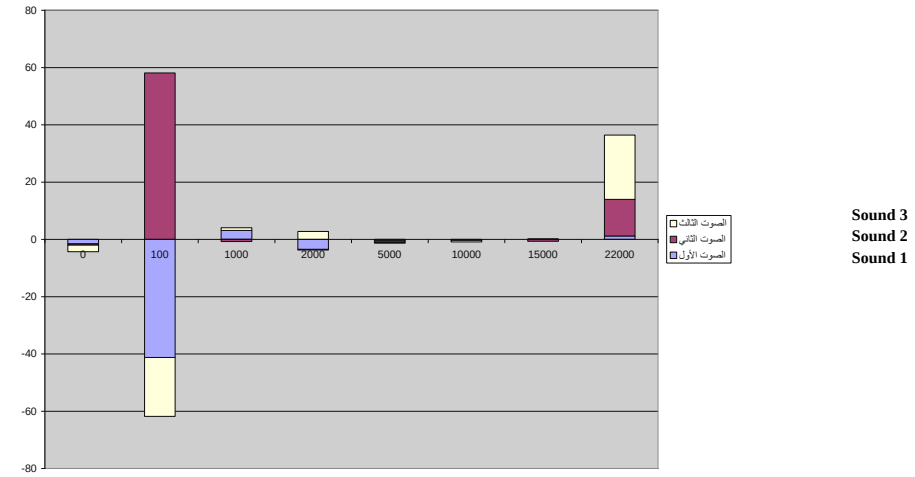


Fig (A10) the coefficient of female speech in frequency domain for age (less than 10 years).

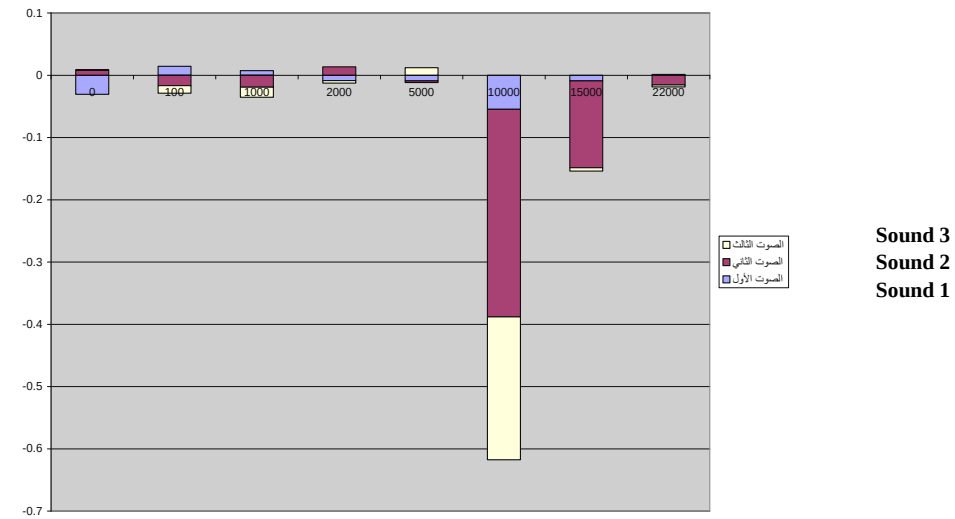


Fig (A11) the coefficient of female speech in time domain for age (less than 10 years).

10-20 year

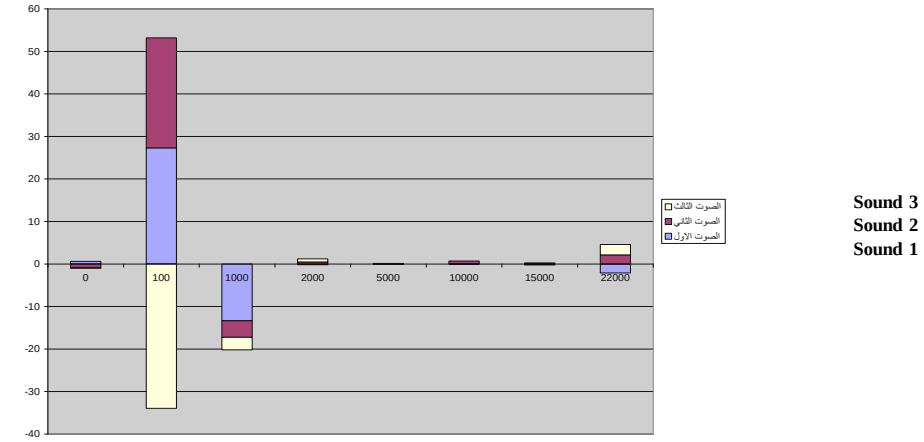


Fig (A12) the coefficient of male speech in frequency domain for age (10-20 years).

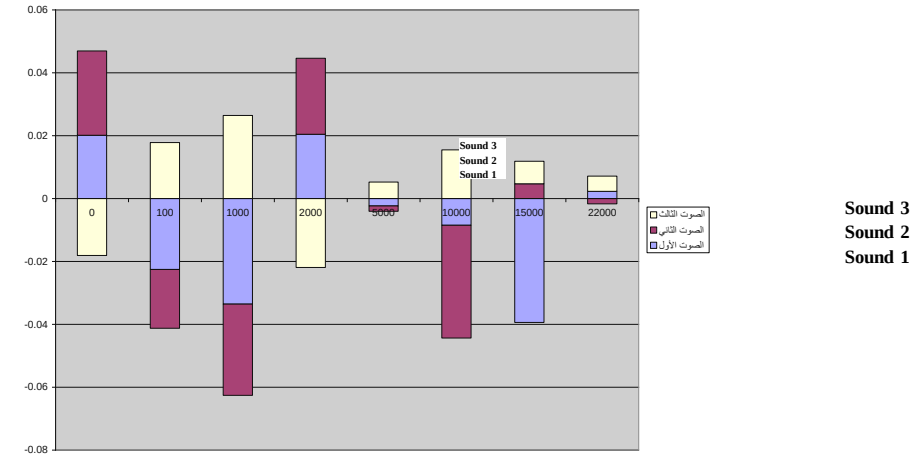


Fig (A13) the coefficient of male speech in time domain for age (10-20 years).

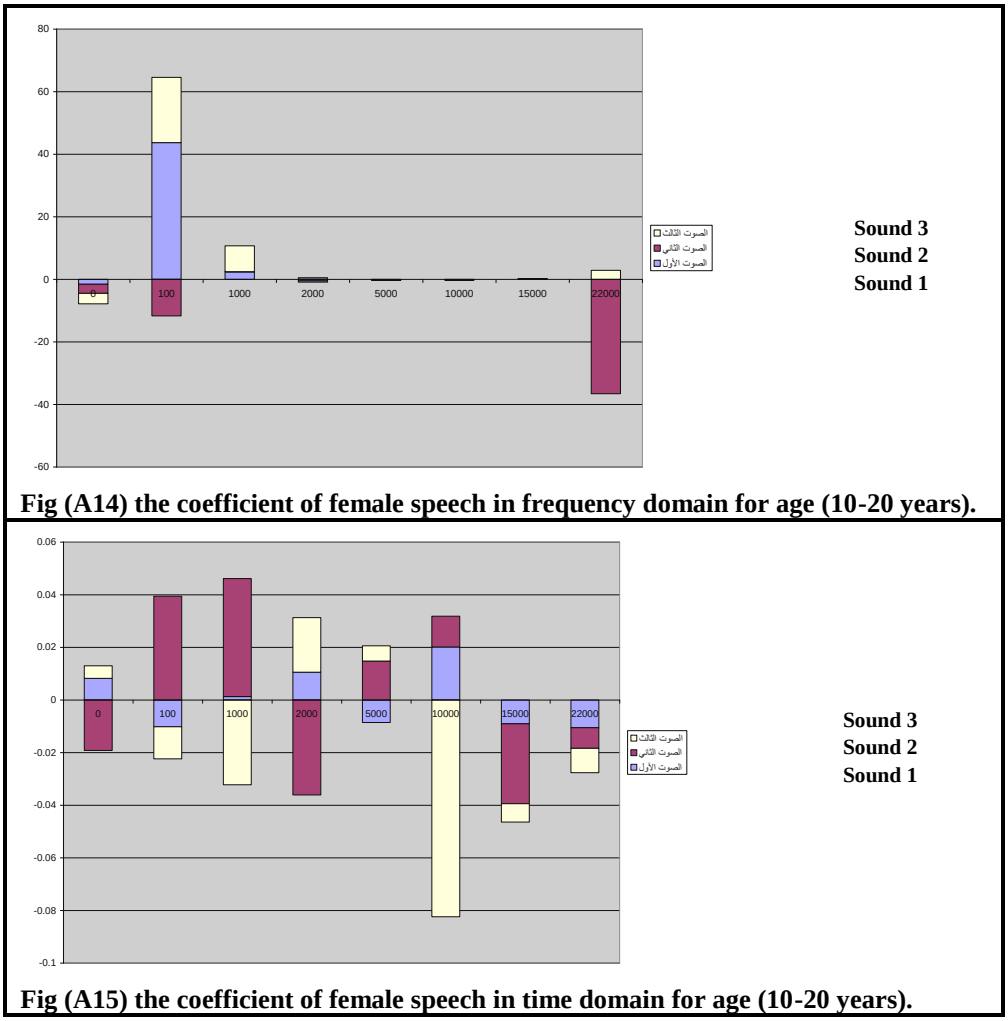


Fig (A14) the coefficient of female speech in frequency domain for age (10-20 years).

Fig (A15) the coefficient of female speech in time domain for age (10-20 years).

20 years and above

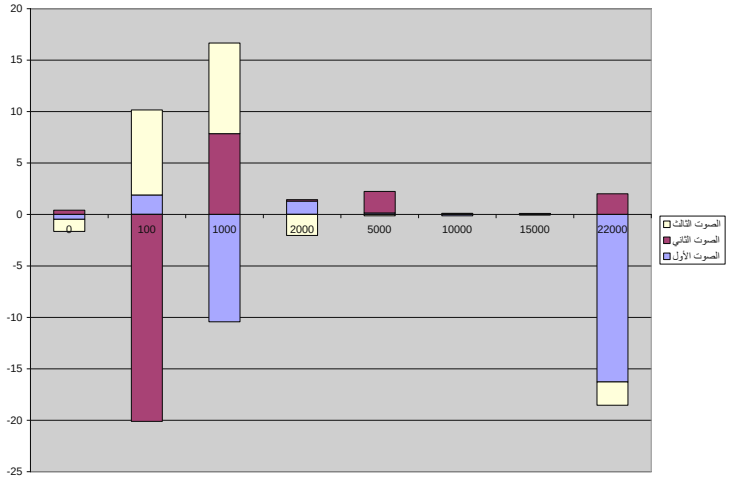


Fig (A16) the coefficient of male speech in frequency domain for age (20 years and above).

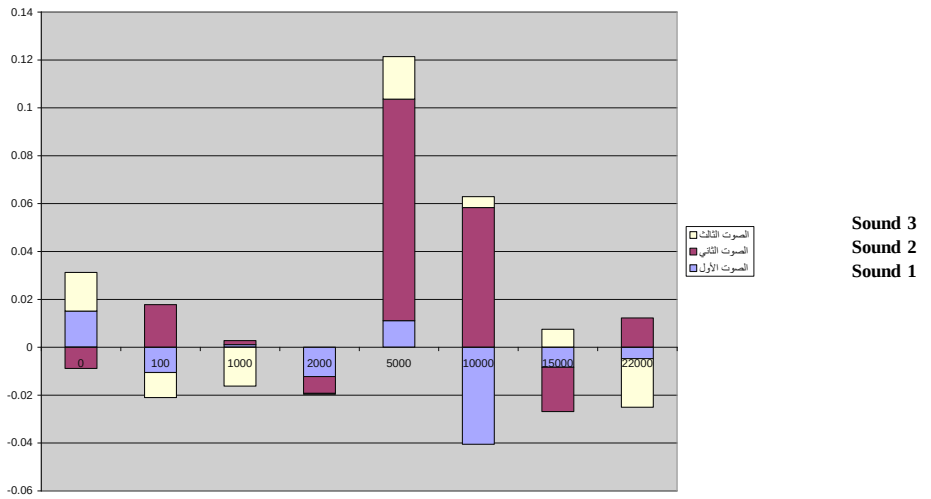


Fig (A17) the coefficient of male speech in time domain for age (20 years and above).

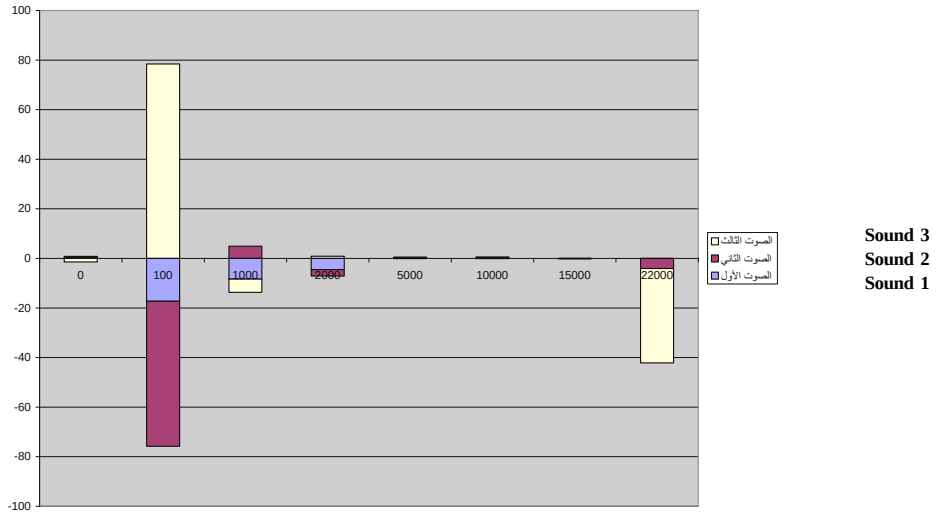


Fig (A18) the coefficient of female speech in frequency domain for age (20 years and above).

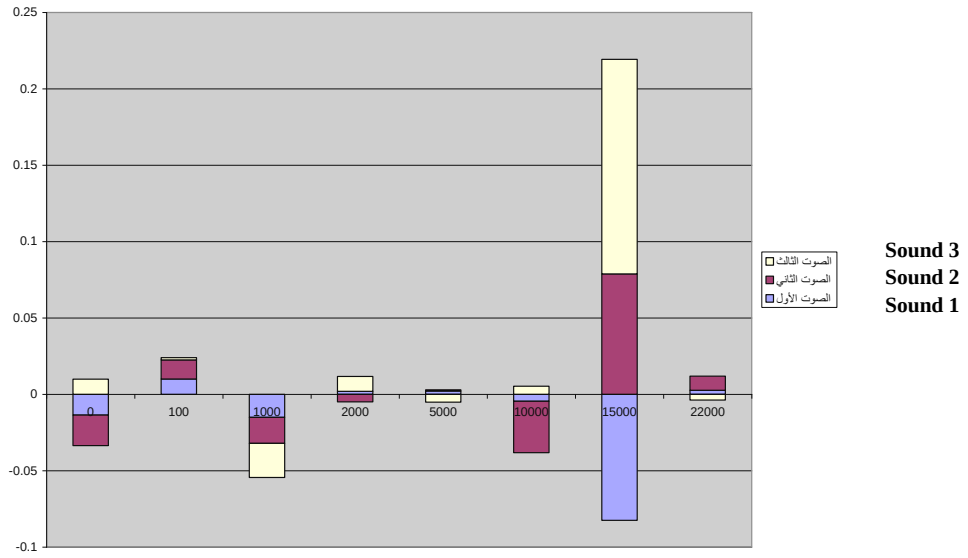


Fig (A19) the coefficient of female speech in time domain for age (20 years and above).

Less than 10 years

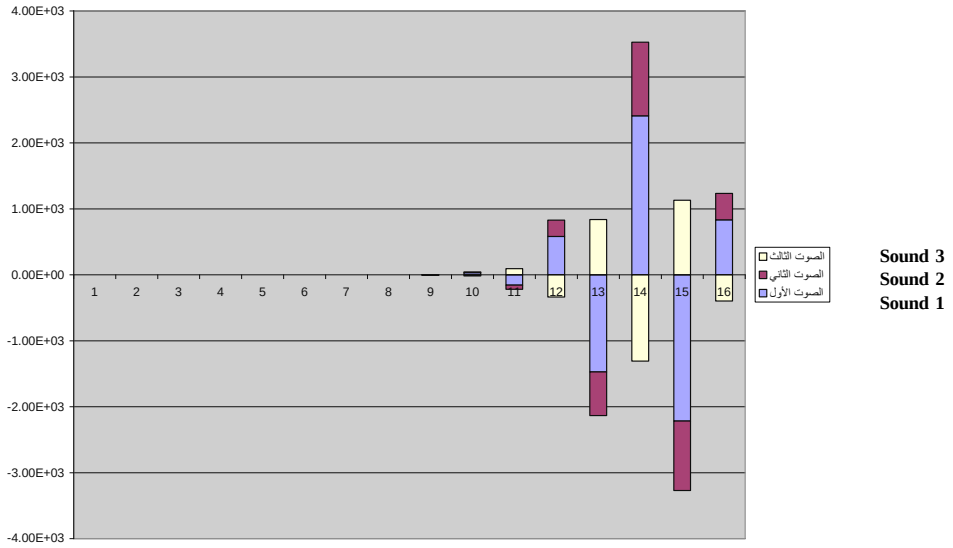


Fig (A20) the coefficient of male polynomial in frequency domain for age (less than 10 years).

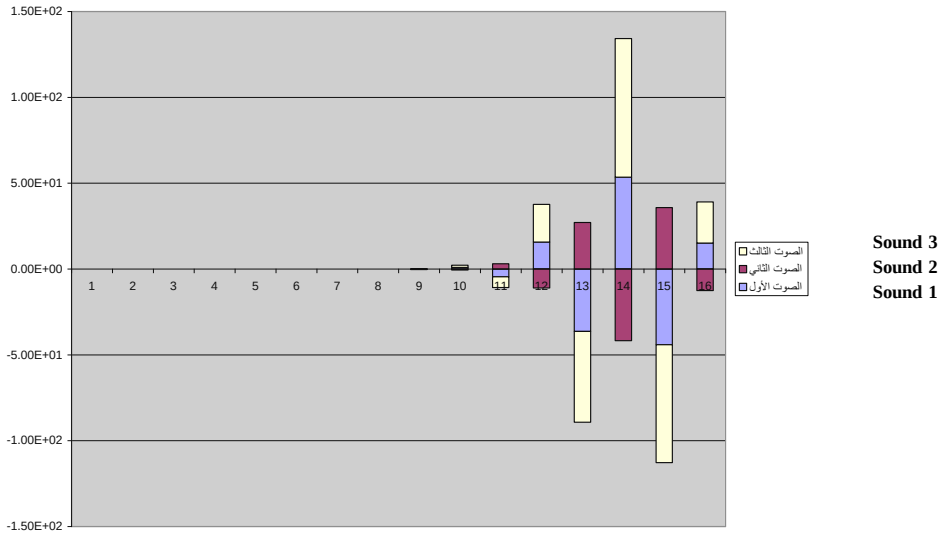


Fig (A21) the coefficient of male polynomial in time domain for age (less than 10 years).

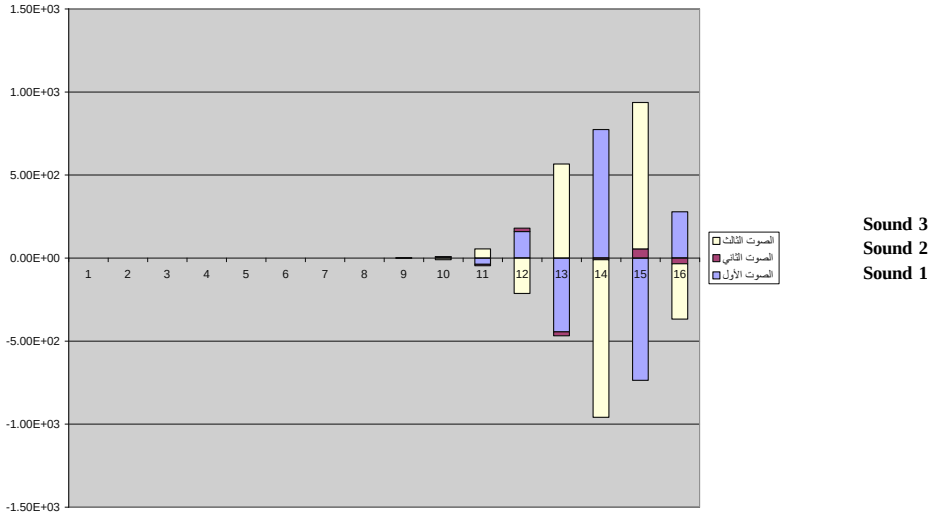


Fig (A22) the coefficient of female polynomial in frequency domain for age (less than 10 years).

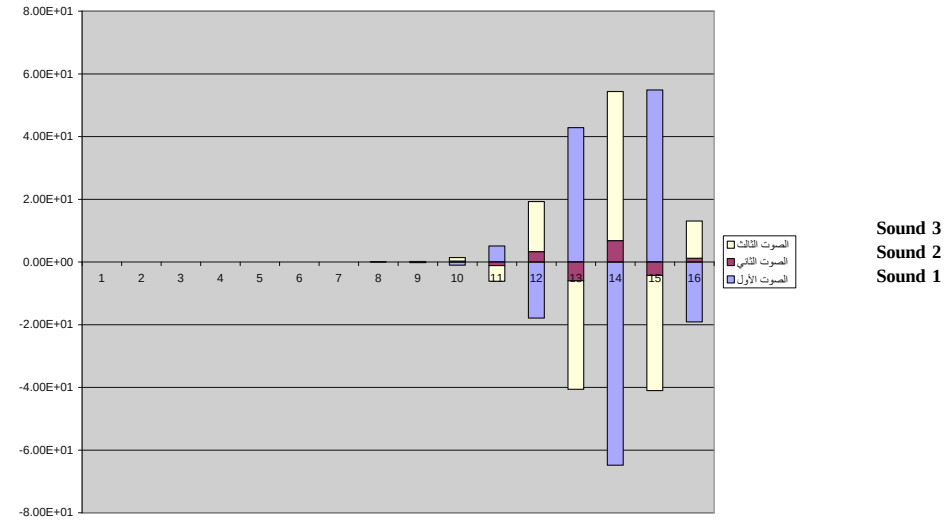


Fig (A23) the coefficient of female polynomial in time domain for age (less than 10 years).

10-20 years

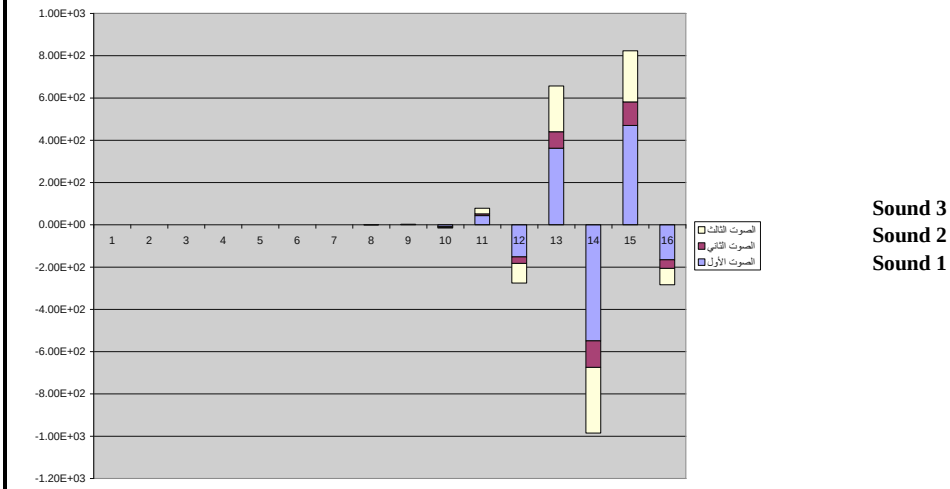


Fig (A24) the coefficient of male polynomial in frequency domain for age (10-20 years).

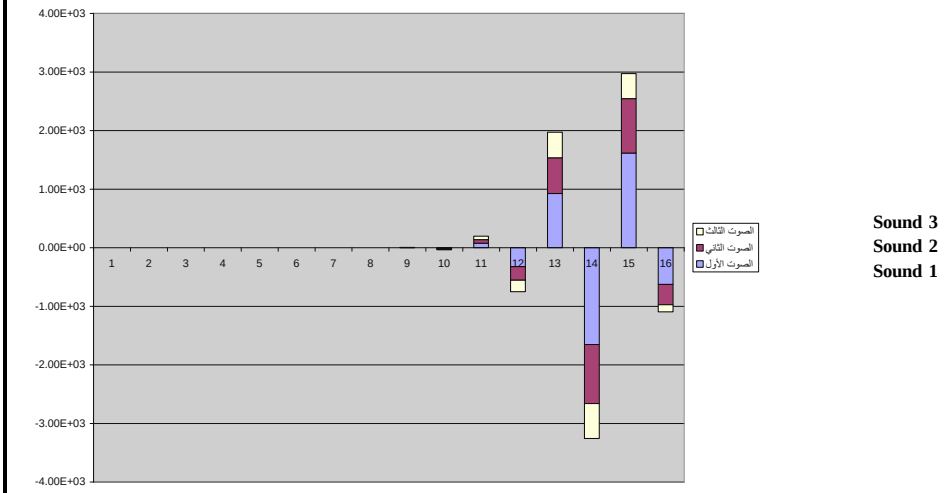


Fig (A25) the coefficient of male polynomial in time domain for age (10-20 years).

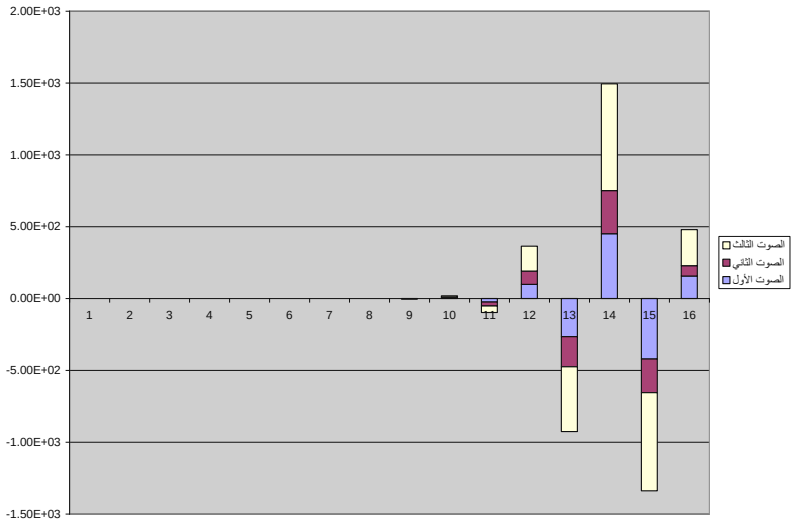


Fig (A26) the coefficient of female polynomial in frequency domain for age (10-20 years).

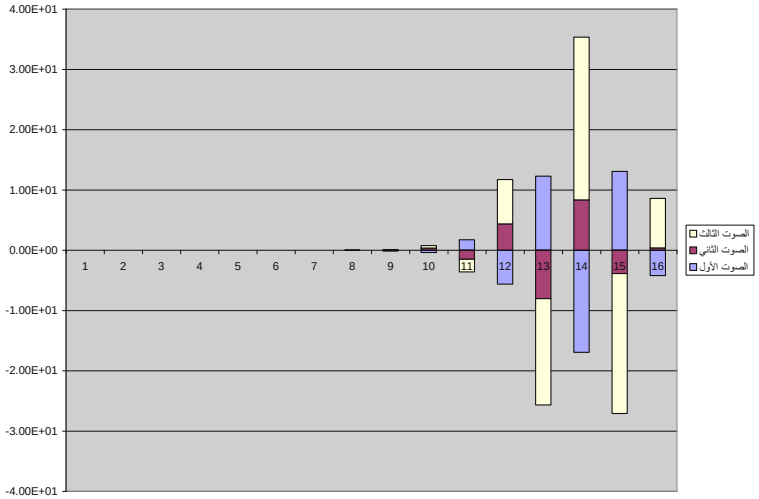


Fig (A27) the coefficient of female polynomial in time domain for age (10-20 years).

20 years and above

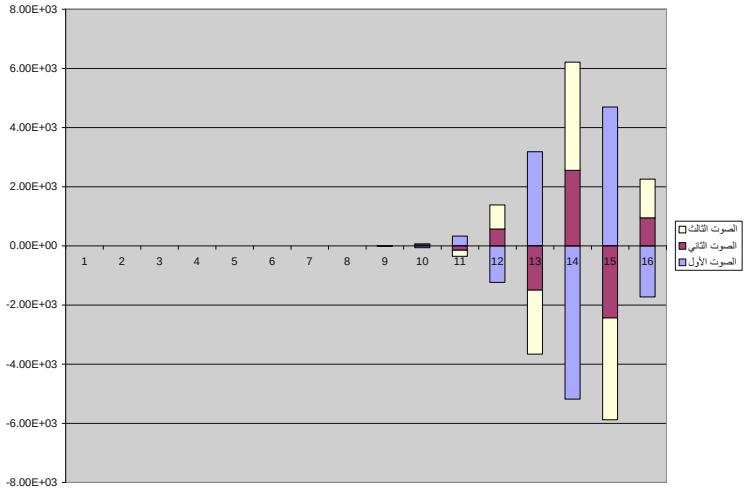


Fig (A28) the coefficient of male polynomial in frequency domain for age (20 years and above).

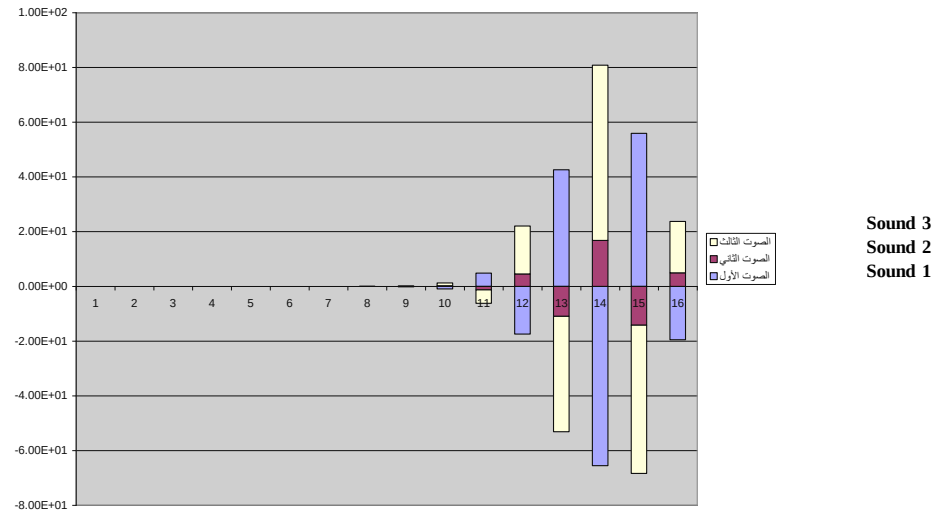


Fig (A29) the coefficient of male polynomial in time domain for age (20 years and above).

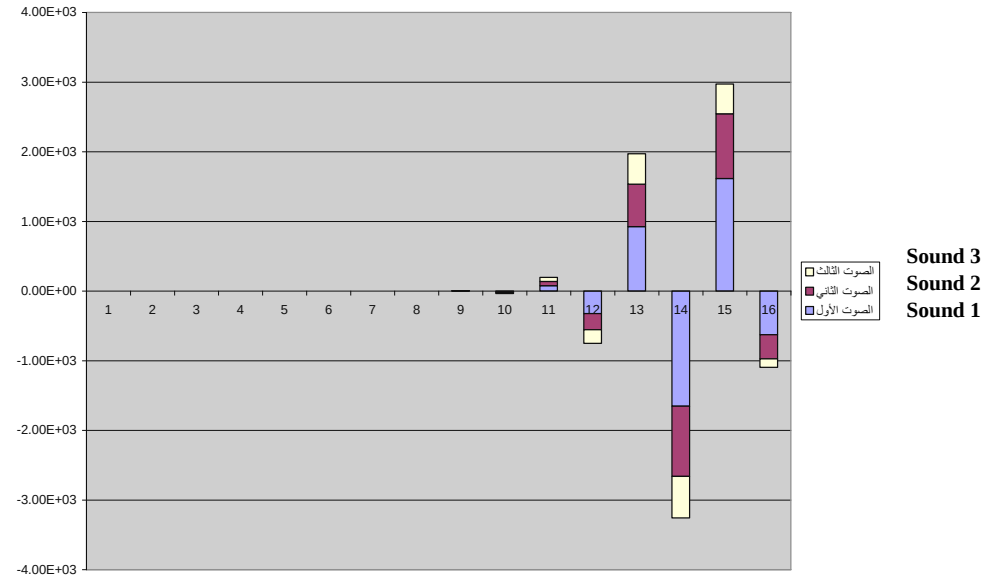


Fig (A30) the coefficient of female polynomial in domain for age (20 years and above).

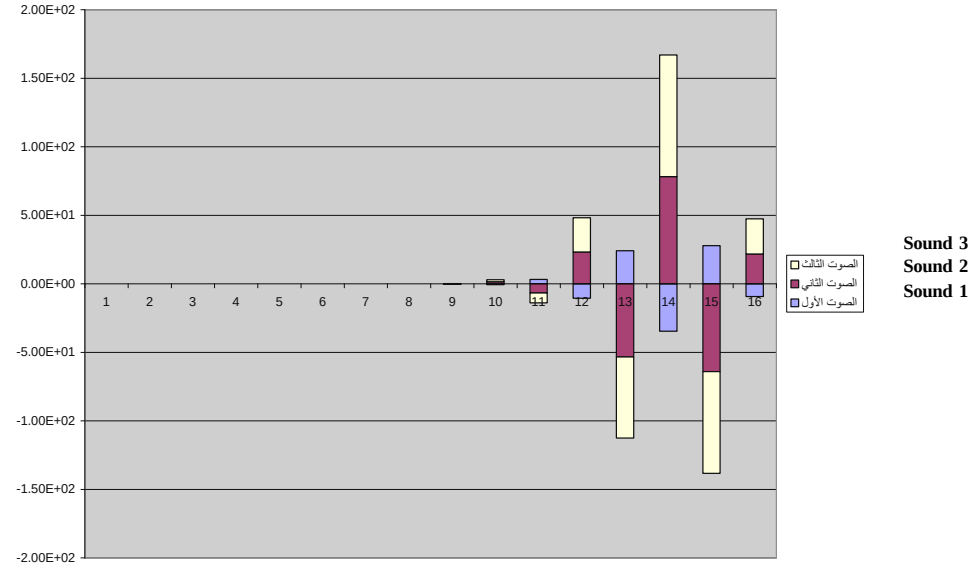


Fig (A31) the coefficient of female polynomial in time domain for age (20 years and above).

Refrences :

[1] Elsom-Cook, M., "Principles of Interactive Multimedia", The McGraw. Hill, 2001.

[2] McEnary, J., "Computer in Music/ What is Sound?" Lectures/ at Materials/Orange Coast College in Costa Mesa, California, 2001.
http://
www.occ.cccd.edu/faculty/jmcenary/sound/sound.html

[3] Minasi, M., "The Complete PC Upgrade and Maintenance Guide", 4th Edition, Sybex-Inc, 1996.

[4] Kivimaki, J., "Very Low Bit Rate Speech Coding Using Speech Recognition, Analysis and Synthesis", M.Sc. Thesis, Department of information Technology, Tampere university of technology, 2000.

[5] Kientzle, T., "Programmer's Guide to Sound", Addison Wesley Developers Press, ISBN 0-201-41972-6, 1997.

[6] Dobrian, C., "Digital Audio", MSP: The Documentation Cycling '74 and IRCAM, Dec, claire trevor school of the arts, University of California, 1997.

[7] Davis, Z., "How Multimedia Works", authorized translation from english language edition, original copyright ©, Ziff-Davis, 1994, translation © Arab scientific publishers 1995.

[8] Stein, J. (y), "Digital Signal Processing", John Wiley & Sons, inc., USA, 2000.

[9] Sherman, L. and et.al., "Digital Audio Conversion", IEEE, ASSP magazine, IEEE Xplor , October 1985 , vol. 2 , issue: 4 , pages 2-25, ISSN: 0740-7467, 1985 .

[10] Shapiro, K., "Graphics and Sound File Formats", Information technology Services, National Library of Canada, ISSN 1201-4338, 1996.

[١١] عبد القادر، إسراء عبد السلام، "كبس الصوت عند الزمن الحقيقي، رسالة ماجستير، كلية علوم الحاسبات والرياضيات ، جامعة الموصل ، ٢٠٠١ .

[12] Skilar, B., "Digital Communication Fundamentals and Application" , 2nd edition , Prentice Hall P T R , Los Angles , 2001.

[13] Hanzo, L., and et.al., "Voice Compression and Communications", IEEE press, Wiley interscience, 2001 .

[14] Gonzales , Rafael C., Woods, Richard E., "Digital Image Processing", 2nd Edition, Prentice-Hall , Inc. , 2002

[١٥] محمد، سجى جاسم، "كبس الصوت بأستخلاص الخواص"، رسالة ماجستير، كلية علوم الحاسبات والرياضيات، جامعة الموصل، ٢٠٠٤.

تصنيف الفئات العمرية باعتماد اشارة الكلام

خليل إبراهيم السيف[١] ، شيماء قيدار الحيالي[٢]

[١] *قسم علوم الحاسبات ، جامعة الموصل ، الموصل ، العراق*

[٢] *مركز الحاسبة ، جامعة الموصل ، الموصل ، العراق*

(تاريخ الاستلام: ٣ / ١١/ ٢٠٠٨ ، تاريخ القبول: ٢٥ / ١٠ / ٢٠٠٩)

الملخص

تم في هذا البحث دراسة العلاقة بين العمر واشارة الكلام التي اعتمدت بثلاثة أساليب:

١-المجال الزمني Time domain .

٢-المجال الترددي Frequency domain .

٣-اعتماد متعدد الحدود Polynomial form .

وقد تبين ان اشارة الكلام بعد اجراء المعالجات الأولية عليها واعتماد حزم (Packets) ذات سعات مختلفة ودراسة إمكانية استخلاص خواص مشتركة لكل فئة عمرية تم اعتمادها في مراحل التمييز.

تبين ان الفئات العمرية عند اعتماد المجال الزمني يمكن ان تعطي مجالات محددة لكل فئة عمرية تعتمد كأساس في توقع عمر المتكلم عند تمييز اشارات الكلام الداخلة .

اتضح من خلال التطبيقات التي اجريت على عينات لفئات عمرية مختلفة ولكلا الجنسين ، ان المجال الترددي (Frequency domain) لتمثيل اشارة الكلام تكون المعاملات فيه متقاربة على العكس مما تم ملاحظته عند اعتماد المجال الزمني (Time domain) .

كذلك تم استنتاج ان الاشارة المسترجعة (عند اعتماد اشارة الكلام بصيغة متعدد الحدود) لا تتأثر عندما تزيد درجة متعدد الحدود عن ١٥ .