

## اعتماد مجموعة عوامل إحصائية لتقطيع أمثل لصوت الكلمة العربية إلى فونيماته الأساسية

أسراء جاسم حرفش

جامعة البصرة/ كلية العلوم / قسم علوم الحاسبات

ISSN-1817-2695

(الاستلام 2008/3/10 ، القبول 2008/7/13)

## المستخلص:

في هذا البحث قمنا بتقطيع صوت الكلمة العربية إلى مجموعة فونيماته الأساسية التي يتألف منها بالاعتماد على ثلاثة عوامل إحصائية مختلفة هي ( المعدل ، الانحراف المعياري ، والتباين) وذلك لأجل تقليل الخطأ في تحديد نقطة القطع الصحيحة إلى اقل ما يمكن . كل عامل من هذه العوامل يتم تطبيقه (على حده) على بيانات صوت الكلمة المدخلة وان نتائج التطبيق هذه تمر بعد ذلك بمجموعة من المعالجات تعطينا نتائج وسطية ، وان النتائج الوسطية المستخرجة من تطبيق كل عامل على البيانات تدمج مع بعضها (لأجل تقليل الخطأ كما ذكرنا في أعلاه) وبذلك تتكون لدينا نتائج موحدة تمر بدورها بمرحلة معالجة لاحقة تكون عندها جاهزة لاستخراج نقاط القطع منها بكفاءة عالية .

**Key Words :** Segmentation , Speech recognition , Statement Segmentation , Word Segmentation , Segments boundaries .

## 1- المقدمة

حدود المقاطع Segments boundaries الصوتية بشكل صحيح ( [2],[3].  
 إن تقطيع الصوت إلى وحداته والتي تعتبر الأساس من ناحية المعالجة والاتصال لمعلوماتي هي إحدى المسائل الأساسية التي تقع تحت تحليل اللغات المنطوقة ، وهي تبقى عملية صعبة بسبب ما تحمله من عدد الإشارات التي تلعب دور في تقطيع الكلام المناسب إلى وحدات صوتية [4] . ففي مهمة التقطيع فإن بعض الأنظمة تعتمد على إشارات وصفات صوتية محددة مركزة على سبيل المثال على السلوك الوقتي أو الاختلافات المحددة بين أصناف الفونيمات [5] . إذا فان مسألة التقطيع الآلي للكلام يمكن ان تشكل تحت مخططين مختلفين ، في النوع الأول فان النسخة الغوية للكلام أو الفونيمات التي تؤلفه هي معروفة ، أما في النوع الثاني فلا توجد معرفة لغوية عن بيانات الصوت المعطاة لذا فان خوارزمية التقطيع هنا تحتاج الإشارات والصفات الصوتية فقط لأجل تحديد عدد وموضع حدود المقاطع في عينات الكلام [6] . إن عملنا في هذا البحث أرتكز على معالجة صوت الكلمة العربية Arabic word signal وتقطيعها إلى فونيماتها الأساسية وإن تركيزنا انصب هنا في استعمال طريقة كفاءة وقلية الخطأ للحصول على قيم واضحة جداً ودقيقة جداً بتطبيق عوامل إحصائية على بيانات الصوت التي تعيد لنا قيماً أمكننا فصلها بسهولة إلى مناطق تمثل

أن عملية تقطيع Segmentation صوت الكلام المنطوق إلى مقاطعه الأساسية هي طريقة واسعة الاستعمال وهي تقود إلى انجازه عالية في عملية تمييز الكلام Speech Recognition . وعلى الرغم من الصعوبة البالغة في الوصول إلى طريقة تقطيع عالية أو تامة الدقة في تحديد نقاط القطع الصحيحة إلا ان البحوث مستمرة وغير متوقفة لأجل الوصول إلى هذا الهدف .  
 إن عملية التقطيع هي عملية أيجاد سلسلة المقاطع الصوتية التي يتألف منها الكلام المنطوق ، وهذه المقاطع قد تكون كلمة أو مقطعاً من كلمة أو فونيم [1] ، أي أن المهتمين في مجال التقطيع يختلفون في توجههم فبعضهم يهتم في تقطيع الجملة المنطوقة Statement Segmentation إلى مجموعة الكلمات الأساسية المتألفة منها وبعضهم يهتم في تقطيع صوت الكلمة Word Segmentation إلى مجموعة مقاطعها الأصلية التي تعتمد عليها اللغة التي يعالجها والبعض الآخر يهتم في تقطيع صوت الكلمة إلى مجموعة فونيماتها الأساسية وهكذا ، إلا ان الشيء المشترك بينهم هو أولاً بذل الجهد العالي في تهيئة أصوات مسجلة بنوعية وجودة عالية من ناحية نقائها من الضوضاء وثانياً معالجة البيانات بطريقة تضمن أن تكون هذه البيانات مهيأة بشكل جيد للتقطيع وثالثاً التركيز في إيجاد طرق قليلة الخطأ في تحديد مناطق القطع الصحيحة ( أي تحديد

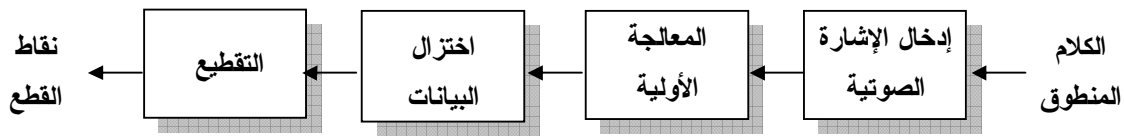
صوت علة وأخرى صوت صحيح من خلال اقتراحنا لمعالجة تقوم بتحديد نقاط القطع من هذه البيانات بكفاءة عالية .

## 2- الصوت اللغوي العربي

تعد اللغة العربية واحدة من اللغات التي اكتسبت أهمية وسيطرة فريدة على مساحات واسعة من البلدان، ويبدل موكب التطور الذي مرت به هذه اللغة منذ بدايتها الأولى حتى اليوم على إنها ميدان خصب للدراسة ، فقد أسهمت دراسات صوتية عديدة للغة العربية منذ القدم إلى الآن في إعطاء نتائج ومواد علمية رائعة عن هذه اللغة [7]. تمتلك اللغة العربية 29 فونيمًا صحيحًا ، وضمن هذه الأصوات صوتان يطلق عليهما أصوات نصف صائته (أو نصف صامتة) وهما (و ، ي) إذ أن الصوت نصف الصائت ينطق بحيث لا يسمح للهواء أن يمر بحرية كما في الصوائت وبنفس الوقت لا يعاق مرورهما كما يحصل في الصوائت ، وتمتلك اللغة العربية (إضافة إلى الأصوات الصحيحة) أيضا 6 فونيمات متحركة ممثلة بثلاثة أصوات طويلة (الألف ، الواو، الياء) وثلاثة أصوات قصيرة (الفتحة، الضمة ، الكسرة) وتعد عادة أصوات جهرة [8] .

## 3- طريقة التقطيع المقترحة

في عملنا هذا هناك أربع مراحل كما موضحة في الشكل (1) ، كل مرحلة تستفيد من النتائج المناسبة من المرحلة السابقة :



شكل (1) وصف النظام المعتمد

### 3-1 إدخال الإشارة الصوتية

سجلت عينات الصوت (حيث كل عينة تمثل كلمة من كلمات اللغة العربية) في جو غرفة واحدة بمعدل نمذجة Sampling Mono . وكانت عينات الصوت المسجلة بصوت الباحث فقط .

### 3-2 المعالجة الأولية للإشارة

تتضمن هذه المرحلة سلسلة من المعالجات التي قمنا بتنفيذها بالتعاقب الآتي:

1- ترشيح الإشارة Filtering وذلك لأجل إزالة الترددات التي هي غير مرغوب فيها في الكلام المنطوق ، لذا استعملنا مرشحاً رقمياً ذا ترشيح واطئ لتنفيذ هذه العملية .

2- تأتي بعد ذلك المعالجة الثانية وهي تجزئة الإشارة Blocking حيث قمنا بتجزئة الإشارة إلى مجموعة من الكتل Frames كل كتلة بطول 128 عينة .

$$X_i = (X_i - 128) / 128 \quad (1)$$

حيث أن  $X_i$  هي العينة  $i$  في الكتلة الواحدة من الإشارة  $X$  .

4- والمعالجة الرابعة هي تطبيق النافذة Windowing لأجل تقليص وحصر نسبة انتشار التوافقيات المرغوب فيها حيث طبقت

$$X_i = X_i * W_i$$

Where

$$W_i = [0.54 - 0.46 \cos(2\pi i / 127)] \dots \dots \dots (2)$$

التحليل التي استعملت هنا هي طريقة تحويل فوريير السريع (FFT) التحليل التي استعملت هنا هي طريقة تحويل فوريير السريع (FFT) حيث أن الاشتقاق الرياضي لخوارزمية FFT يمكن كتابته كالآتي [10]:

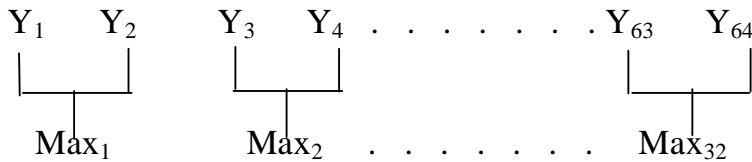
$$A_f = \sum_{t=0}^{N-1} X_t e^{-2\pi j f t / N} = \sum_{t=0}^{(N/2)-1} X_{2t} e^{-2\pi j f (2t) / N} + \sum_{t=0}^{(N/2)-1} X_{2t+1} e^{-2\pi j f (2t+1) / N} \dots \dots \dots (3)$$

6- وهي مرحلة تنعيم المعاملات Parameters Smoothing لوحظ في الكثير من التجارب أن هناك نسبة من الضوضاء الملحقة بالإشارة التي تتغير ساعاتها بشكل عشوائي وسريع من نقطة إلى نقطة ضمن الإشارة ، لذا فإن التقنية المستعملة لأجل تقليص نسبة هذه الضوضاء تعرف بعملية التنعيم Smoothing التي يمكن تطبيقها على معاملات الإشارة الناتجة من مرحلة التحليل السابقة ، وان معادلة التنعيم هي:

$$Y_i = \begin{cases} Y = (Y_{i-1} + Y_{i+1}) / 2 & \text{IF } Y_i < Y \\ Y_i & \text{Otherwise} \end{cases} \dots \dots \dots (4)$$

سوف نشير من الآن فصاعداً إلى مجموعة المعاملات الناتجة من مرحلة المعالجة الأولية للإشارة الموضحة أعلاه بـ FFT64 التي هي مصفوفة بطول N من الأسطر وهو عدد كتل الإشارة الداخلة و64 عمود الذي هو طول الكتلة الواحدة .

بمقدار النصف ، ويتم حفظ هذه البيانات المختزلة الجديدة في مصفوفة جديدة خاصة بها ، وفيما يلي توضيح لخطوات معالجة الاختزال:  
1- لكل كتلة من المصفوفة FFT64 و بطول 64 تقلص إلى كتلة بطول 32 عينة وكالاتي:



حيث  $1 \leq i \leq 128$  -5 استخراج المعاملات Coefficients Calculation وذلك لأجل تحليل الإشارة واستخراج المعاملات المهمة منها ، وطريقة

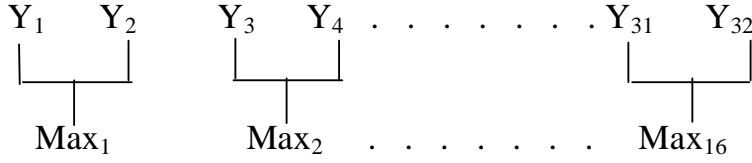
ومما يذكر هنا أن تحويلات فوريير تتطلب العمل على مدى يمتد من 0 إلى N لذا فإن الفترة من N/2 إلى N هي الجزء المناظر للجزء من 0 إلى N/2 (حيث N هنا طول كل كتلة وهو 128). لذا فيكون طول كل كتلة بعد عملية التحليل في بحثنا هذا هو  $2^6 = 64$  قيمة من المعاملات بسبب عملية التناظر هذه ممتدة من القيمة التي تسلسلها 2 وانتهاء بالقيمة 65 لأن القيمة الأولى تمثل الـ DC (التردد صفر) وقيمتها عادةً صفر .

3-3 اختزال البيانات

إن لهذه المرحلة من المعالجة دوراً مهماً في تحديد سير تنفيذ المرحلة الأخيرة التي هي مرحلة التقطيع ، إذ إن المقصود من اختزال البيانات هنا هي مجموعة من الخطوات المتعاقبة كل خطوة منها تقوم بعملية اختزال البيانات الداخلة إليها ينتج عنها مجموعة بيانات جديدة ذات حجم اقل من حجم البيانات المعالجة الداخلة



2- لكل كتلة من المصفوفة FFT32 وبطول 32 تقلص إلى كتلة بطول 16 عينة بنفس الطريقة أعلاه وكالاتي :



أي لكل قيمتين متجاورتين نأخذ أكبر قيمة بينهما ، وتكرر هذه العملية على كل الكتل فتكون لدينا مصفوفة جديدة FFT32 طول الكتلة فيها هو 32 عينة و N من الأسطر .

فتكون لدينا مصفوفة جديدة FFT8 طول الكتلة فيها هو 8 عينة و N من الأسطر .

وخالصة ناتج هذه المرحلة هي المصفوفات , FFT8 , FFT16 , FFT32 التي ستكون الإدخال إلى المرحلة القادمة بالإضافة إلى المصفوفة الأساسية FFT64 .

فتكون لدينا مصفوفة جديدة FFT16 طول الكتلة فيها هو 16 عينة و N من الأسطر .

3- في الخطوة الثالثة فأن لكل كتلة من المصفوفة FFT16 وبطول 16 تقلص إلى كتلة بطول 8 عينة وبفهم الطريقة أعلاه،

### 3-4 التقطيع

3- للمصفوفة FFT16 تطبيق العوامل الإحصائية الثلاثة بنفس الطريقة الموضحة في خطوة 1 فتكون لدينا أيضا مصفوفة جديدة أسمها Test16 مكونة من N من الأسطر و 3 أعمدة .

4- للمصفوفة FFT8 تطبيق العوامل الإحصائية الثلاثة بنفس الطريقة الموضحة في خطوة 1 فتكون لدينا أيضا مصفوفة جديدة أسمها Test8 مكونة من N من الأسطر و 3 أعمدة . وفيما يلي مثال يوضح المعالجة بالعوامل الإحصائية، فالإشارة الصوتية التي أخذت كمثال هنا هي إشارة الكلمة " شاش " المسجلة بصوت الباحثة والموضحة في الشكل (2) والتي هي بطول 46 كتلة ( أي N=46 ) . فبعد أن تمر هذه الإشارة بسلسلة المراحل الموضحة في البنود السابقة في أعلاه إلى أن تصل إلى مرحلة المعالجة بالعوامل الإحصائية يكون إخراج هذه المرحلة لهذه الإشارة الصوتية النتائج الموضحة في الجداول (1) و (2) و (3) و (4) على التوالي وحسب الخطوات أعلاه الموضحة في هذا البند .

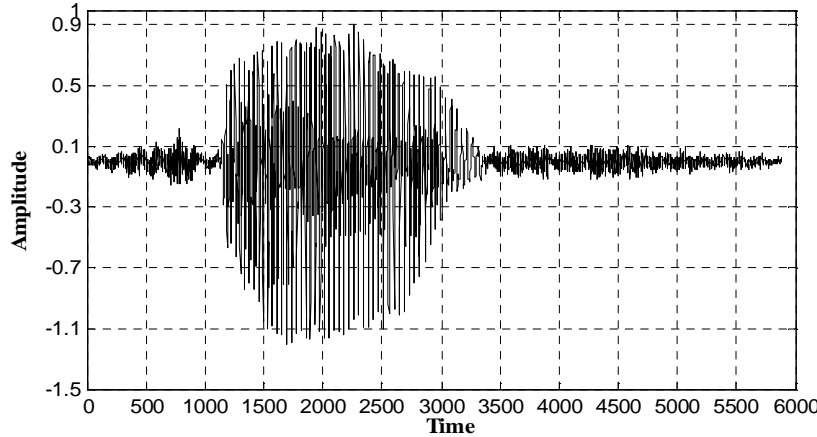
تتقسم هذه المرحلة إلى ثلاث معالجات تنفذ بالتعاقب كما سيتم توضيحه في أدناه.

### 1-3-4 تطبيق العوامل الإحصائية

يتم هنا تطبيق العوامل الإحصائية التي هي المعدل والانحراف المعياري والتباين Covariance على التوالي على كل مصفوفة من المصفوفات الأربعة , FFT16 , FFT8 , FFT64 وكالاتي :

1- للمصفوفة FFT64 تطبيق العوامل الإحصائية الثلاثة على كل كتلة (سطر) فيها حيث إن ناتج تطبيق كل عامل على مجموعة بيانات الكتلة الواحدة تعيد قيمة واحدة ، فتطبيق المعدل (mean) على عناصر الكتلة K يعيد قيمة توضع في مصفوفة جديدة أسمها Test64 في السطر K والعمود 1 ، ثم يطبق الانحراف المعياري (Std) على نفس الكتلة وتعيد قيمة واحدة توضع في السطر K والعمود 2 ، ثم على نفس الكتلة يطبق التباين (Cov) والنتيجة توضع في السطر K العمود 3 ، وهكذا إلى نهاية الكتل في المصفوفة FFT64 فتكون المصفوفة الجديدة Test64 بطول N من الأسطر و 3 أعمدة .

2- للمصفوفة FFT32 تطبيق العوامل الإحصائية الثلاثة بنفس الطريقة الموضحة في خطوة 1 فتكون لدينا أيضا مصفوفة جديدة أسمها Test32 مكونة من N من الأسطر و 3 أعمدة .



شكل (2) رسم للإشارة الصوتية التي تمثل كلمة "شاش"

	<i>mean</i>	<i>Std</i>	<i>Cov</i>
1.	0.10951	0.071932	0.0051742
2.	0.18921	0.14829	0.021989
3.	0.2408	0.18322	0.033571
4.	0.28515	0.21844	0.047715
5.	0.33936	0.26828	0.071974
6.	0.42277	0.35619	0.12687
7.	0.43683	0.35499	0.12602
8.	0.29726	0.21186	0.044886
9.	0.72067	0.82932	0.68778
10.	1.4124	1.6916	2.8616
11.	2.015	2.2616	5.1148
12.	2.3981	2.4557	6.0303
13.	2.6169	2.5883	6.6995
14.	2.6669	2.6717	7.1379
15.	2.615	2.5738	6.6243
16.	2.5489	2.4435	5.9705
17.	2.5009	2.3185	5.3754
18.	2.4673	2.2322	4.9827
19.	2.3963	2.1879	4.7871
20.	2.2197	2.0698	4.2839
21.	2.1345	2.004	4.0161
22.	1.7976	1.6685	2.7839
23.	1.5341	1.4634	2.1414
24.	1.1146	1.1717	1.3729
25.	0.73982	0.96493	0.9311
26.	0.44502	0.64005	0.40966
27.	0.2869	0.3312	0.10969
28.	0.2875	0.20155	0.040622
29.	0.31471	0.28862	0.083304
30.	0.29563	0.27389	0.075016
31.	0.2829	0.25397	0.064503
32.	0.26551	0.20317	0.041278
33.	0.28713	0.23789	0.056591
34.	0.31757	0.28298	0.080078
35.	0.29378	0.27234	0.074167
36.	0.26656	0.22151	0.049065
37.	0.28482	0.30058	0.090349
38.	0.24247	0.20203	0.040817
39.	0.24096	0.16704	0.027901
40.	0.2541	0.16191	0.026214
41.	0.21337	0.12356	0.015267
42.	0.18825	0.11752	0.013811
43.	0.19305	0.15966	0.02549
44.	0.17487	0.12383	0.015335
45.	0.15038	0.10946	0.011981
46.	0.076375	0.044465	0.0019772

جدول (2) المصفوفة Test32 الناتجة .

	<i>mean</i>	<i>Std</i>	<i>Cov</i>
1.	0.089916	0.078099	0.0060995
2.	0.16042	0.16154	0.026094
3.	0.20837	0.19214	0.036916
4.	0.25056	0.237	0.056171
5.	0.28966	0.28059	0.078731
6.	0.34519	0.35083	0.12308
7.	0.35939	0.36478	0.13307
8.	0.24164	0.19788	0.039157
9.	0.53603	0.90576	0.8204
10.	1.0974	1.8583	3.4531
11.	1.5263	2.3836	5.6818
12.	1.7548	2.5197	6.3488
13.	1.8865	2.5805	6.6591
14.	1.9296	2.6709	7.1336
15.	1.8932	2.6035	6.7782
16.	1.899	2.5272	6.3866
17.	1.8271	2.4129	5.8221
18.	1.7926	2.3447	5.4975
19.	1.7638	2.19	4.7961
20.	1.6676	2.0186	4.0747
21.	1.6013	1.9443	3.7801
22.	1.41	1.6842	2.8365
23.	1.1928	1.5087	2.2762
24.	0.83817	1.2183	1.4843
25.	0.56646	1.0152	1.0306
26.	0.33731	0.65904	0.43433
27.	0.22751	0.33123	0.10971
28.	0.2375	0.18851	0.035537
29.	0.25764	0.2965	0.08791
30.	0.24083	0.29115	0.084769
31.	0.23394	0.26724	0.071418
32.	0.22564	0.21347	0.045571
33.	0.23624	0.24131	0.05823
34.	0.26742	0.30237	0.091427
35.	0.25027	0.29642	0.087863
36.	0.2312	0.25205	0.063529
37.	0.23023	0.3048	0.092904
38.	0.20712	0.21708	0.047124
39.	0.20229	0.16737	0.028012
40.	0.2123	0.15241	0.023229
41.	0.18486	0.11627	0.013519
42.	0.15868	0.11134	0.012397
43.	0.16612	0.17092	0.029212
44.	0.15513	0.1307	0.017081
45.	0.13111	0.12139	0.014736
46.	0.058594	0.046725	0.0021832

جدول (1) المصفوفة Test64 الناتجة .

	mean	Std	Cov
1.	0.17715	0.062683	0.0039292
2.	0.33283	0.13589	0.018467
3.	0.38862	0.15517	0.024078
4.	0.461	0.18153	0.032951
5.	0.56935	0.23129	0.053496
6.	0.738	0.32197	0.10367
7.	0.76824	0.31409	0.098653
8.	0.49443	0.19297	0.037238
9.	1.1152	0.61294	0.3757
10.	2.1511	1.2374	1.5311
11.	3.1738	1.7808	3.1714
12.	3.7727	1.9432	3.7761
13.	3.952	1.9523	3.8116
14.	4.0941	2.0572	4.2322
15.	4.0341	2.0205	4.0824
16.	4.0013	1.9116	3.6542
17.	3.8506	1.7752	3.1513
18.	3.6364	1.6532	2.7331
19.	3.6024	1.6746	2.8044
20.	3.3564	1.6032	2.5702
21.	3.1379	1.5147	2.2942
22.	2.4209	1.1706	1.3702
23.	2.2502	1.1322	1.282
24.	1.6688	0.90638	0.82153
25.	1.2109	0.7777	0.60482
26.	0.70975	0.49322	0.24327
27.	0.45374	0.26257	0.068941
28.	0.42524	0.16203	0.026254
29.	0.50908	0.23062	0.053187
30.	0.46179	0.20955	0.043913
31.	0.4282	0.20113	0.040454
32.	0.3957	0.16124	0.026
33.	0.45792	0.19218	0.036933
34.	0.55742	0.2463	0.060662
35.	0.52706	0.25428	0.064656
36.	0.46362	0.19762	0.039053
37.	0.49085	0.24567	0.060353
38.	0.40182	0.16811	0.02826
39.	0.38168	0.13734	0.018861
40.	0.37226	0.12893	0.016624
41.	0.31032	0.10533	0.011094
42.	0.30758	0.10909	0.0119
43.	0.35909	0.15642	0.024467
44.	0.29858	0.11062	0.012238
45.	0.22366	0.085826	0.0073662
46.	0.11496	0.038774	0.0015034

جدول (4) المصفوفة Test8 الناتجة .

	mean	Std	Cov
1.	0.13077	0.06392	0.0040857
2.	0.23371	0.13889	0.019292
3.	0.29245	0.16561	0.027425
4.	0.34823	0.19471	0.037912
5.	0.42932	0.25299	0.064005
6.	0.5341	0.33594	0.11285
7.	0.55022	0.32611	0.10635
8.	0.36401	0.19946	0.039784
9.	0.92077	0.74236	0.55109
10.	1.8607	1.5282	2.3353
11.	2.6299	2.0949	4.3885
12.	3.0728	2.2803	5.1997
13.	3.073	2.1542	4.6408
14.	3.2224	2.2943	5.2637
15.	3.2549	2.2966	5.2743
16.	3.2769	2.2256	4.9535
17.	3.1894	2.1145	4.4712
18.	3.0282	2.0103	4.0412
19.	2.9838	2.0085	4.0341
20.	2.8781	1.9562	3.8266
21.	2.7649	1.8793	3.5317
22.	2.181	1.4761	2.179
23.	1.9273	1.3525	1.8293
24.	1.3933	1.0661	1.1366
25.	0.89633	0.83929	0.7044
26.	0.51836	0.52137	0.27182
27.	0.32423	0.26687	0.071219
28.	0.33297	0.17424	0.03036
29.	0.3833	0.24623	0.060628
30.	0.35246	0.22595	0.051054
31.	0.31292	0.20864	0.043529
32.	0.30476	0.17066	0.029126
33.	0.3533	0.20571	0.042318
34.	0.38872	0.25155	0.063277
35.	0.3709	0.25912	0.067145
36.	0.33132	0.20344	0.041387
37.	0.33177	0.24718	0.061099
38.	0.27861	0.17093	0.029217
39.	0.28341	0.14498	0.021019
40.	0.29861	0.14295	0.020434
41.	0.24584	0.11279	0.012721
42.	0.23201	0.11368	0.012924
43.	0.25399	0.15861	0.025156
44.	0.21245	0.1139	0.012972
45.	0.18035	0.099075	0.0098158
46.	0.091871	0.040375	0.0016302

جدول (3) المصفوفة Test16 الناتجة .

## 3-4-2 معالجة البيانات

يتم هنا معالجة البيانات التي تم الحصول عليها من المعالجة في البند 1-2-4 ، حيث أن كل مصفوفة من المصفوفات Test8 , Test64 , Test32 , Test16 تعالج على حده بالشكل الآتي:

1- يتم تقسيم قيمة كل عمود في المصفوفة على معدل ذلك العمود .

2- يضاف عمود رابع لكل مصفوفة يتضمن معدل القيم للأعمدة 1 , 2 , 3 .

والجدول رقم (5) بين التغيرات الحاصلة على مصفوفة Test64 بعد تنفيذ الخطوتين أعلاه

	<i>mean</i>	<i>Std</i>	<i>Cov</i>	<i>Average</i>
1.	0.13098	0.087352	0.0036529	0.073994
2.	0.23367	0.18067	0.015627	0.14333
3.	0.30352	0.2149	0.022109	0.18018
4.	0.36497	0.26508	0.03364	0.22123
5.	0.42193	0.31383	0.047151	0.26097
6.	0.50282	0.3924	0.073712	0.32298
7.	0.5235	0.408	0.079692	0.33707
8.	0.35198	0.22133	0.02345	0.19892
9.	0.7808	1.0131	0.49132	0.76173
10.	1.5985	2.0784	2.068	1.915
11.	2.2232	2.6661	3.4027	2.764
12.	2.5561	2.8182	3.8022	3.0588
13.	2.7479	2.8862	3.988	3.2074
14.	2.8107	2.9873	4.2722	3.3567
15.	2.7578	2.912	4.0594	3.243
16.	2.7662	2.8266	3.8249	3.1392
17.	2.6614	2.6988	3.4868	2.949
18.	2.6111	2.6225	3.2923	2.842
19.	2.5692	2.4495	2.8723	2.6303
20.	2.4291	2.2577	2.4403	2.3757
21.	2.3325	2.1746	2.2638	2.257
22.	2.0539	1.8837	1.6987	1.8788
23.	1.7374	1.6875	1.3632	1.596
24.	1.2209	1.3627	0.88893	1.1575
25.	0.82513	1.1355	0.61721	0.85927
26.	0.49134	0.73712	0.26011	0.49619
27.	0.3314	0.37047	0.065704	0.25586
28.	0.34596	0.21085	0.021283	0.1927
29.	0.37528	0.33162	0.052648	0.25319
30.	0.35081	0.32565	0.050767	0.24241
31.	0.34076	0.2989	0.042771	0.22748
32.	0.32868	0.23877	0.027292	0.19825
33.	0.34411	0.2699	0.034873	0.21629
34.	0.38953	0.33819	0.054754	0.26083
35.	0.36455	0.33154	0.05262	0.24957
36.	0.33678	0.28191	0.038047	0.21891
37.	0.33536	0.34091	0.055639	0.24397
38.	0.30169	0.2428	0.028222	0.1909
39.	0.29467	0.1872	0.016776	0.16621
40.	0.30924	0.17047	0.013911	0.16454
41.	0.26928	0.13005	0.0080961	0.13581
42.	0.23113	0.12453	0.0074243	0.12103
43.	0.24197	0.19116	0.017495	0.15021
44.	0.22598	0.14618	0.01023	0.12746
45.	0.19099	0.13577	0.0088249	0.11186
46.	0.085351	0.052261	0.0013075	0.046307

جدول (5) قيم المصفوفة Test64 التي تم معالجتها .

بعد تنفيذ الخطوتين في أعلاه على التوالي يتم بناء المصفوفة 2- ثم تطرح قيمة المعدل هذه من كل قيمة من قيم نفس العمود النهائية المسماة Final وكالاتي:

1- لكل مصفوفة من مصفوفات Test8 , Test16 , Test32 , Test64 يتم إيجاد معدل العمود الرابع منها.

Final ذات N من الأسطر و 4 أعمدة .

### 3-4-3 استخراج نقاط القطع

من المصفوفة النهائية Final بالإمكان الآن تحديد نقاط القطع واستخراجها وكالاتي:

1- يضاف عمود خامس إلى المصفوفة Final يتضمن معدل القيم للأعمدة 1 , 2 , 3 , 4 كما هو موضح في جدول (6) والخاص بالإشارة في الشكل (2) . وأن عملية المعدل هذه

التي قمنا بها في هذه الخطوة هي لأجل تقليل نسبة الخطأ إلى أقل مايمكن في عملية تحديد نقاط القطع والموضحة في خطوة 2 التالية ، وكذلك تسهيل عملية تحديد نقاط القطع .

2- بالتالي يكون العمود الخامس (عمود المعدل) كما هو موضح في الجدول (6) عبارة عن قيم سالبة وأخرى موجبة

Count=0;

For i=1 : N-1

If (Final(i,5)< 0)and( Final(i+1,5) >=0) OR  
(Final(i,5)>= 0)and( Final(i+1,5)<0)

Count=Count+1 ;

SegPoint(count)= i ;

End;

فالقيم السالبة المتوالية هي منطقة صوت صحيح والقيم الموجبة المتوالية هي منطقة صوت علة لذلك يتم البحث عن نقاط الانقلاب من سالب إلى موجب فيأخذ رقم موقع القيمة السالبة كنقطة قطع بين صوت صحيح وآخر علة، وكذلك البحث عن نقاط الانقلاب من موجب إلى سالب فيأخذ رقم موقع القيمة الموجبة كنقطة قطع بين صوت علة وآخر صحيح ، وبذلك يتم تحديد مقاطع الصوت الداخلة بشكل مضبوط ودقيق كما موضح في أدناه:

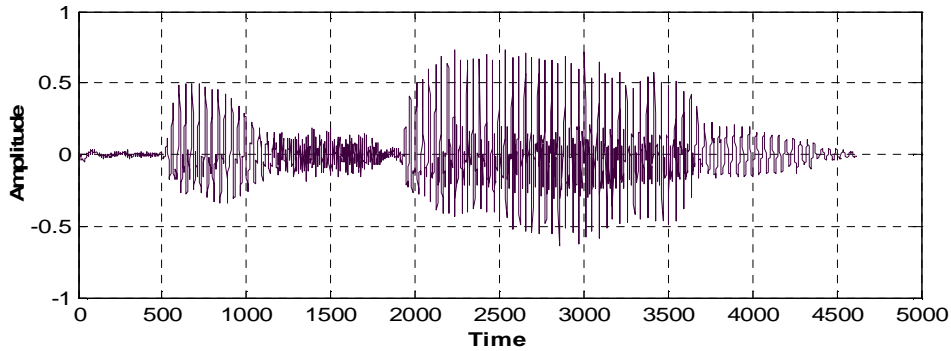
	عمود1 Test64	عمود2 Test32	عمود3 Test16	عمود4 Test8	المعدل Average	
1.	-0.92576	-0.93083	-0.93283	-0.92519	-3.7146	ش
2.	-0.8911	-0.8998	-0.89972	-0.886	-3.5766	
3.	-0.85534	-0.86689	-0.86775	-0.854	-3.444	
4.	-0.81693	-0.83406	-0.83598	-0.8213	-3.3083	
5.	-0.77786	-0.7956	-0.79236	-0.77613	-3.1419	
6.	-0.72732	-0.73848	-0.73116	-0.70223	-2.8992	
7.	-0.69501	-0.70764	-0.70109	-0.66448	-2.7682	
8.	-0.74792	-0.75375	-0.74907	-0.71897	-2.9697	
9.	-0.49298	-0.51383	-0.5097	-0.50911	-2.0256	
10.	0.21112	0.12476	0.14497	0.07671	0.55756	
11.	0.98768	0.89867	0.97001	0.91446	3.7708	
12.	1.5234	1.4887	1.5915	1.5514	6.155	
13.	1.8655	1.919	1.8038	1.8999	7.4883	
14.	2.1112	2.205	2.0418	2.191	8.549	
15.	2.1773	2.2662	2.1674	2.294	8.9048	
16.	2.1584	2.1917	2.1772	2.2393	8.7666	
17.	2.0538	2.0597	2.0837	2.0719	8.2691	
18.	1.948	1.9302	1.9357	1.859	7.6729	
19.	1.7893	1.8235	1.8539	1.7662	7.2329	
20.	1.5826	1.6626	1.7596	1.6316	6.6364	
21.	1.4199	1.526	1.6415	1.4678	6.0552	
22.	1.1495	1.2032	1.2369	1.0527	4.6423	
23.	0.87286	0.88711	0.92586	0.79982	3.4857	
24.	0.5153	0.51598	0.54162	0.46717	2.0401	
25.	0.18741	0.17581	0.1723	0.17624	0.71176	
26.	-0.15808	-0.16499	-0.19135	-0.16279	-0.67721	
27.	-0.45099	-0.45479	-0.48172	-0.45032	-1.8378	
28.	-0.62903	-0.63154	-0.65054	-0.62971	-2.5408	
29.	-0.6878	-0.69382	-0.70833	-0.68777	-2.7777	
30.	-0.72258	-0.73215	-0.74732	-0.72927	-2.9313	
31.	-0.74743	-0.75856	-0.77731	-0.75672	-3.04	
32.	-0.77447	-0.7871	-0.80343	-0.78669	-3.1517	
33.	-0.77897	-0.78916	-0.80017	-0.78469	-3.153	
34.	-0.75895	-0.77351	-0.78086	-0.7543	-3.0676	
35.	-0.75457	-0.77272	-0.77174	-0.74008	-3.0391	
36.	-0.76771	-0.78967	-0.78818	-0.759	-3.1046	
37.	-0.76175	-0.77538	-0.7845	-0.74964	-3.0713	
38.	-0.7853	-0.80002	-0.81083	-0.78033	-3.1765	
39.	-0.80942	-0.82065	-0.82985	-0.80729	-3.2672	
40.	-0.82232	-0.8297	-0.83761	-0.82436	-3.314	
41.	-0.84314	-0.8502	-0.8567	-0.84709	-3.3971	
42.	-0.86093	-0.86639	-0.86808	-0.85771	-3.4531	
43.	-0.85524	-0.86433	-0.8594	-0.84302	-3.422	
44.	-0.86377	-0.87456	-0.87227	-0.85632	-3.4669	
45.	-0.87583	-0.88728	-0.88701	-0.87889	-3.529	
46.	-0.91464	-0.92075	-0.92099	-0.91579	-3.6722	

جدول (6) المصفوفة النهائية Final



وكما نلاحظ من جدول (6) فإن أول نقطة قطع هي الكتلة التي رقمها 9 والتي هي نقطة انتقال من الصوت الصحيح "ش" (التي حدوده مضللة في الجدول) إلى صوت العلة "الألف" وذلك لان القيمة في الموقع Final(9,5) سالبة والقيمة التالية لها في الموقع Final(10,5) هي موجبة ، أما نقطة القطع التالية فهي الكتلة التي رقمها 25 فهي نقطة انتقال من صوت العلة "الألف" إلى الصوت الصحيح الثاني "ش" لان القيمة في الموقع

Final(25,5) هي موجبة والقيمة التالية لها في الموقع Final(25,5) هي سالبة .  
وفيما يلي أمثلة أخرى لأصوات أخرى توضح طريقتنا الموصوفة في هذا البحث:  
مثال (1): الصوت المدخل هو صوت الكلمة "هُشام" :

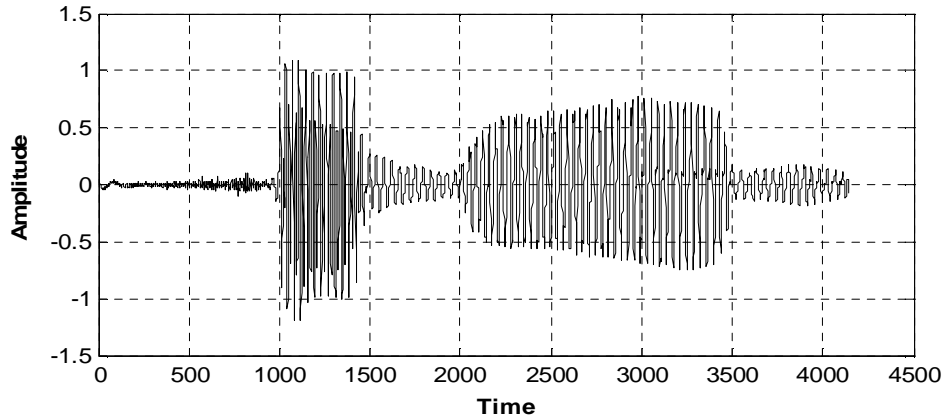


شكل (3) رسم للإشارة الصوتية التي تمثل كلمة "هُشام"

جدول (7) للمصفوفة النهائية Final لإشارة الصوت "هُشام"

	عمود1 Test64	عمود2 Test32	عمود3 Test16	عمود4 Test8	المعدل Average
1.	-0.90148	-0.9055	-0.89446	-0.88043	-3.5819
2.	-0.91116	-0.91527	-0.9086	-0.8978	-3.6328
3.	-0.9262	-0.93027	-0.92747	-0.92114	-3.7051
4.	-0.66828	-0.70723	-0.70258	-0.69626	-2.7743
5.	-0.12312	-0.25052	-0.23946	-0.22283	-0.83594
6.	0.31071	0.20343	0.16018	0.079798	0.75412
7.	0.50684	0.42483	0.34344	0.22029	1.4954
8.	0.4579	0.39186	0.36331	0.35734	1.5704
9.	0.12205	0.091392	0.05919	0.094754	0.36738
10.	-0.19122	-0.19526	-0.21602	-0.18367	-0.78618
11.	-0.34879	-0.33762	-0.34756	-0.34291	-1.3769
12.	-0.43652	-0.43611	-0.44224	-0.461	-1.7759
13.	-0.50779	-0.51455	-0.52171	-0.53466	-2.0787
14.	-0.5965	-0.59764	-0.6116	-0.61883	-2.4246
15.	-0.42356	-0.4898	-0.47023	-0.46959	-1.8532
16.	0.15233	-0.056807	0.023368	0.04389	0.16278
17.	0.54737	0.34737	0.43308	0.32074	1.6486
18.	0.71866	0.6177	0.70551	0.60149	2.6434
19.	0.72798	0.69951	0.80446	0.71475	2.9467
20.	0.72104	0.74473	0.86385	0.75323	3.0828
21.	0.74017	0.80071	0.90372	0.79197	3.2366
22.	0.76536	0.85928	0.92635	0.88088	3.4319
23.	0.79831	0.91542	0.92198	0.92131	3.557
24.	0.84038	0.95855	0.89896	0.90219	3.6001
25.	0.71448	0.81154	0.75675	0.74439	3.0272
26.	0.56089	0.65951	0.59714	0.57795	2.3955
27.	0.45601	0.56614	0.49777	0.48594	2.0059
28.	0.38164	0.49368	0.4212	0.41608	1.7126
29.	0.19573	0.27268	0.22137	0.26451	0.95429
30.	0.012651	0.061254	0.0302	0.1146	0.21871
31.	-0.14842	-0.1133	-0.13775	-0.020687	-0.42016
32.	-0.27441	-0.2465	-0.26627	-0.13842	-0.9256
33.	-0.38964	-0.37239	-0.3897	-0.26956	-1.4213
34.	-0.53028	-0.5151	-0.52054	-0.42495	-1.9909
35.	-0.66725	-0.65687	-0.65726	-0.58876	-2.5701
36.	-0.79127	-0.7865	-0.78603	-0.74356	-3.1074
37.	-0.89461	-0.89232	-0.89234	-0.87106	-3.5503

مثال (2): الصوت المدخل هو صوت الكلمة " سبيل " :



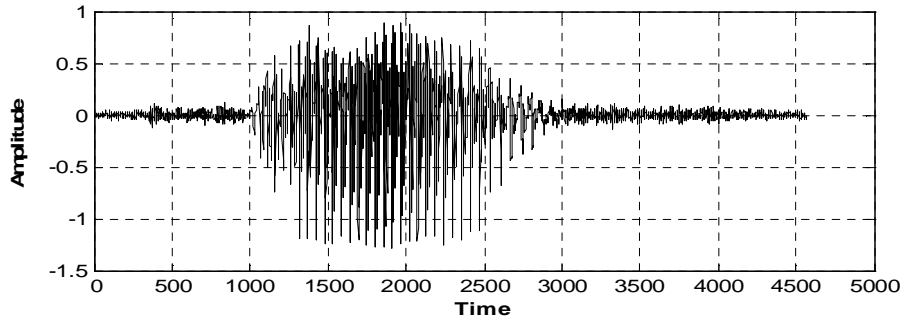
شكل (4) رسم للإشارة الصوتية التي تمثل كلمة " سبيل "

جدول (8) المصفوفة النهائية لإشارة الصوت " سبيل "

	عمود1 Test64	عمود2 Test32	عمود3 Test16	عمود4 Test8	المعدل Average
1.	-0.93129	-0.93139	-0.93085	-0.92676	-3.7203
2.	-0.93179	-0.93374	-0.93618	-0.93604	-3.7377
3.	-0.93509	-0.9379	-0.94295	-0.94773	-3.7637
4.	-0.93025	-0.9334	-0.94017	-0.94797	-3.7518
5.	-0.91824	-0.91992	-0.93032	-0.94069	-3.7092
6.	-0.89754	-0.89877	-0.91031	-0.91998	-3.6266
7.	-0.83074	-0.83385	-0.84384	-0.8512	-3.3596
8.	-0.24644	-0.23875	-0.21167	-0.1963	-0.89316
9.	1.0309	1.0824	1.1855	1.2539	4.5526
10.	1.6038	1.8679	1.7371	1.7693	6.9781
11.	1.7109	2.1408	1.7578	1.7173	7.3268
12.	1.0348	1.3587	1.0492	0.90436	4.347
13.	-0.29117	-0.48562	-0.31635	-0.22829	-1.3214
14.	-0.18029	-0.060864	-0.15138	-0.17254	-0.56507
15.	-0.44218	-0.36703	-0.41964	-0.41175	-1.6406
16.	-0.41954	-0.41194	-0.42752	-0.3987	-1.6577
17.	0.19513	0.2602	0.23288	0.1778	0.86602
18.	0.10557	0.016441	0.049813	0.14669	0.31851
19.	0.25268	0.10616	0.19347	0.30652	0.85883
20.	0.32777	0.17258	0.27099	0.38595	1.1573
21.	0.40064	0.24905	0.35659	0.46557	1.4718
22.	0.46753	0.32082	0.43958	0.53848	1.7664
23.	0.54516	0.39221	0.52168	0.59134	2.0504
24.	0.60932	0.44437	0.58225	0.60189	2.2378
25.	0.66084	0.4753	0.60087	0.55057	2.2876
26.	0.69241	0.53907	0.64386	0.51621	2.3915
27.	0.63218	0.55406	0.60902	0.4669	2.2622
28.	0.2238	0.19499	0.22271	0.13221	0.7737
29.	-0.24007	-0.23165	-0.22677	-0.26215	-0.96064
30.	-0.46635	-0.44503	-0.44859	-0.45565	-1.8156
31.	-0.57919	-0.54978	-0.55517	-0.56415	-2.2483
32.	-0.66126	-0.6363	-0.64169	-0.66208	-2.6013
33.	-0.78405	-0.7771	-0.78667	-0.80397	-3.1518

- من خلال جدول (7) الناتج عن صوت الكلمة "مُشام" فأن مناطق القطع التي تنتج هي : 5 ، 9 ، 15 ، 30 على التوالي ، ووفقاً لنقاط القطع الناتجة هذه فأن حدود المقاطع الصوتية التي سوف تستخرج هي:
- 1- الكتلة من 1 إلى 5 تمثل الفونيم " هـ "
  - 2- الكتلة من 6 إلى 9 تمثل الفونيم " الضمة "
  - 3- الكتلة من 10 إلى 15 تمثل الفونيم " ش "
  - 4- الكتلة من 16 إلى 30 تمثل الفونيم " الإلف "

- 5- الكتلة من 31 إلى الكتلة الأخيرة 37 تمثل الفونيم "م" ومن خلال جدول (8) الناتج عن صوت الكلمة "سبيل" فإن مناطق القطع التي تنتج هي : 8 ، 12 ، 16 ، 28 على التوالي ، ووفقاً لنقاط القطع الناتجة هذه فإن حدود المقاطع الصوتية التي سوف تستخرج هي:
- 1- الكتلة من 1 إلى 8 تمثل الفونيم "س"
- 2- الكتلة من 9 إلى 12 تمثل الفونيم "الفتحة"
- 3- الكتلة من 13 إلى 16 تمثل الفونيم "ب"
- 4- الكتلة من 17 إلى 28 تمثل الفونيم "ي"
- 5- الكتلة من 29 إلى الكتلة الأخيرة 33 تمثل الفونيم "ل"
- مثال (3): الصوت المدخل هو صوت الكلمة "شاش" كما في شكل (5) لكن مسجل بصوت شخص آخر ، والجدول (9) يبين ناتج معالجة هذا الصوت :



شكل (5) رسم للإشارة الصوتية التي تمثل كلمة "شاش" بصوت شخص آخر

جدول (9) المصفوفة النهائية Final لإشارة الصوت "شاش" بصوت شخص آخر

	عمود 641 Test	عمود 322 Test	عمود 163 Test	عمود 4 Test	المعدل Average
1.	-0.94116	-0.92963	-0.92217	-0.91406	-3.707
2.	-0.92411	-0.91785	-0.91139	-0.90235	-3.6557
3.	-0.90908	-0.9042	-0.89729	-0.88965	-3.6002
4.	-0.88593	-0.88093	-0.87178	-0.86405	-3.5027
5.	-0.87373	-0.87246	-0.8688	-0.86257	-3.4776
6.	-0.85269	-0.85348	-0.8544	-0.853	-3.4136
7.	-0.82991	-0.82932	-0.8325	-0.82903	-3.3208
8.	-0.62788	-0.62827	-0.663	-0.66779	-2.5869
9.	-0.16979	-0.19855	-0.31058	-0.32433	-1.0032
10.	0.2505	0.24528	0.099387	0.13054	0.72571
11.	0.86187	0.97134	0.80124	0.94062	3.5751
12.	1.2434	1.376	1.3165	1.4845	5.4204
13.	1.417	1.4941	1.5514	1.6116	6.0741
14.	1.7004	1.7458	1.8616	1.8462	7.154
15.	1.99	2.0342	2.1598	2.1097	8.2937
16.	2.1476	2.1573	2.3047	2.1762	8.7858
17.	2.1036	2.0552	2.1852	2.0362	8.3801
18.	1.9461	1.8585	1.9261	1.8313	7.5621
19.	1.6243	1.5188	1.5571	1.4669	6.167
20.	1.1906	1.1618	1.1527	1.1435	4.6485
21.	0.64646	0.62863	0.57493	0.5896	2.4396
22.	0.20549	0.16384	0.10495	0.09816	0.57244
23.	-0.16965	-0.20622	-0.23987	-0.23969	-0.85544
24.	-0.44796	-0.46834	-0.48496	-0.48308	-1.8843
25.	-0.61283	-0.62442	-0.6415	-0.64142	-2.5202
26.	-0.70744	-0.71387	-0.71985	-0.72178	-2.8629
27.	-0.765	-0.76924	-0.7696	-0.77001	-3.0739
28.	-0.80404	-0.8068	-0.80618	-0.80812	-3.2251
29.	-0.81429	-0.81609	-0.8178	-0.81558	-3.2638
30.	-0.80951	-0.81267	-0.81439	-0.79922	-3.2358
31.	-0.82378	-0.82751	-0.82563	-0.80542	-3.2823
32.	-0.84406	-0.84474	-0.84678	-0.82975	-3.3653
33.	-0.85842	-0.85531	-0.85844	-0.84393	-3.4161
34.	-0.87107	-0.86803	-0.8679	-0.85399	-3.461
35.	-0.88331	-0.88362	-0.87851	-0.86592	-3.5114
36.	-0.90162	-0.89923	0.89227	-0.88032	-3.5734

ش

ا

ش

من خلال جدول (9) الناتج عن صوت الكلمة "شاش" فأُن

مناطق القطع التي تنتج هي : 9 ، 22 على التوالي.

#### 4- الاستنتاج Conclusion

أن الطريقة المقترحة في بحثنا هذا والمعتمدة على مجموعة من العوامل الإحصائية في معالجة البيانات هي لأجل الوصول إلى بيانات يمكن أن تستعمل وبشكل كفوء جداً في الحصول على نقاط قطع دقيقة من الإشارة الداخلة وقد استنتجنا الأتي:

- 1- قمنا بتجربة تطبيق الكثير من العوامل الإحصائية الأخرى (معامل الارتباط RMS و corrcoeff) ووجدنا إن انسب عوامل هي العوامل التي اعتمدت في هذا العمل وذلك لأنها أكثر العوامل التي أعطتنا نتائج مؤثرة وفعالة جداً في عملية التقطيع ، فمن بين 50 صوت كلمة مسجلة أعطت هذه الخوارزمية تقطيع صحيح لـ 36 صوت
- 2- أن اعتمادنا على ثلاثة عوامل إحصائية معاً كان له الأثر الكبير في رفع كفاءة عملية التقطيع مقارنة بالنتائج التي يمكن أن نحصل عليها بتطبيق كل عامل على حدة في عملية التقطيع ، إذ قمنا بتجربة الاعتماد على تطبيق كل عامل إحصائي وحده فوجدنا ان نسبت النجاح اقل مما هو مدون في نقطة 1 وهي دائماً محصورة بين 50 % إلى 60 % .
- 3- أن بناء المصفوفة Final كان له دور فعال جداً في استخراج نقاط القطع بسهولة ودقة عالية .

#### المصادر Reference

- [1] S. Cassidy , "Speech Recognition" , Speech Hearing and Language Research Center , University of Macquarie , Sydney, Australia , Steve ,[2001] .
- [2] A. Jordi and B. Antonio , " Towards Phone Segmentation for Concatenative Speech Synthesis" , Dept. of Signal Theory and Comunication ,TALP Research Center, Technical University of Catalonia (UPC), Barcelona (Spain) , [2004].
- [3] A. Valery , "Adaptive Algorithms for Pitch-synchronous Speech Signal Segmentation" , Accenture Technology Labs ,161 N. Clark St., Chicago, IL 60601, USA ,[2003] .
- [4] A. Mettouchi , "Only Prosody? Perception of speech segmentation in Kabyle and Hebrew" , Laboratory of Linguistique of Nantes, University of Nantes ,[2007]
- [5] K. Demuynck and T. Laureys , "A Comparison of Different Approaches to Automatic Speech Segmentation " , K.U.Leuven ESAT/PSI , Kasteelpark Arenberg 10 ,B-3001 Leuven, Belgium ,[2002]
- [6] M. Sharma, " "BLIND" SPEECH SEGMENTATION: AUTOMATIC SEGMENTATION OF SPEECH WITHOUT LINGUISTIC KNOWLEDGE" , SpeakeZ, Inc.,a subsidiary of T-NETIX, Inc.371 Hoes Lane, Piscataway, NJ 08854,msharma@tnetix.com , [1996].
- [7] العاني ، سلمان حسن ، " التشكيل الصوتي في اللغة العربية: فونولوجيا العربية" ، كندا ، [1966] .
- [8] حسان ، تمام ، " منهاج البحث في اللغة " ، كلية دار العلوم ، القاهرة.
- [9] P. Lynn , "Digital Signal Processing with Computer Application" , Second Edition , John Wiley & Sons ,[1997] .
- [10] J. Holmes , "Speech Synthesis and Recognition" , Speech Technology Consultant , UK ,[1988] .

## Using group of statistical operator toward Optimal segmentation of Arabic word signal to its basic phonemes

**Esra'a Jassim Harfash**  
**University of Basrah**  
**Department of Computer Science**

### **Abstract:**

In this research, we are segment the sound Arabic signal to its basic group of phonemes depending on three statistical operators that are( mean ,standard deviation ,covariance) to reduce the error in determining the correct points segment . Each operator is implemented on the data of the sound entered word and then the results are manipulated by some processes that give intermediate results , these results (from each implementing of operator) merges together and give union results that are passed to postprocessing in order to prepare to find the points segment in high efficient.