مجلة جامعة تشرين للبحوث والدراسات العلمية _ سلسلة العلوم الهندسية المجلد (37) العدد (6) 2015
Tishreen University Journal for Research and Scientific Studies - Engineering Sciences Series Vol. (37) No. (6) 2015

توليد وتحليل المضلعات الصوتية للهجات سورية اعتماداً على قاعدة بيانات صوتية مُنجَزة

الدكتور جعفر الخير•
رنيم كناج••

(تاريخ الإيداع 13 / 8 / 2015. قُبِل للنشر في 10/ 11 / 2015)

□ ملخّص □

تشكل قواعد البيانات الصوتية حجر الأساس في بناء نظم النطق الآلي والتعرف الآلي على الكلام وعلى المتكلم باختلاف اللغات واللهجات، وتكون عناصر قاعدة البيانات الصوتية عادةً ملفات صوتية سبق أن سجلت لأصوات أشخاص باللغة أو اللهجة المطلوبة، وكلما ذخرت قاعدة البيانات الصوتية بعناصر شاملة كلما أسهم ذلك في إنتاج أنظمة تخاطب مع الآلة ذات أداء متميز.

ونظراً لعدم توفر قاعدة بيانات صوتية للهجات السورية؛ تمَّ في البحث بإنشاء قاعدة بيانات للصوتيات السورية، ضمت ستة عشر متطوع ومتطوعة من مناطق سورية مختلفة اللهجة سُجلت أصواتهم في ظروف تسجيل مختلفة؛ وذلك لدراسة تأثير تنوع اللهجات واختلاف الجنس وظروف التسجيل على مساحات المضلعات الصوتية.

استثمر هذا البحث قاعدة البيانات الصوتية المُنجزة في مجال توليد وتحليل مضلعات الصوتيات السورية، وبالتالي تمَّ الحصول 64 مضلع صوتي تعود إلى 16 متحدث،حيث أن المضلع الصوتي هو مضلع هندسي رؤوسه تمثل قيم ترددات توافقيات المجرى الصوتي الممتد من الحبال الصوتية إلى الشفاه وتسمى النغمات (Formants)، ومساحة المضلع الصوتي الناتج تمثل مساحة المجال الصوتي.

الكلمات المفتاحية: قاعدة البيانات الصوتية،المضلعات الصوتية، المجال الصوتي، خوارزمية MFCC.

• أستاذ مساعد، قسم هندسة الحاسبات والتحكم الآلي، كلية الهندسة الميكانيكية والكهربائية، جامعة تشرين، اللاذقية سورية.
••طالبة دراسات عليا(ماجستير) ،قسم هندسة الحاسبات والتحكم الآلي،كلية الهندسة الميكانيكية والكهربائية، جامعة تشرين،اللاذقية سورية.

مجلة جامعة تشرين للبحوث والدراسات العلمية _ سلسلة العلوم الهندسية المجلد (37) العدد (6) 2015
Tishreen University Journal for Research and Scientific Studies - Engineering Sciences Series Vol. (37) No. (6) 2015

Generation and Analysis of Vowel Polygons for Syrian Dialects Using a Created Speech Database

Dr. Jaffar Alkhier •
Raneem Knaj ••

(Received 13 / 8 / 2015. Accepted 10 / 11 / 2015)

□ ABSTRACT □

Speech databases form the main foundation in the construction of automatic utterance, speaker recognition and speech recognition systems in different languages and dialects. The elements of the speech database are audio files recorded for people's voices in the required language or dialect. The more the speech database is enriched with comprehensive elements the more it contributes to produce systems that communicate with the excellent performed machine. According to the lack of speech databases for the Syrian dialects, the research did one. The created database contained sixteen voluntaries from different Syrian dialects. Voluntaries' voices were recorded in different recording conditions that is for studying the effect of variety of dialects, gender and the conditions of recording on the vowel polygons. This research invested the created speech database in the field of generating and analyzing of vowel polygons, as the vowel polygon is a geometric polygon where its vertices represent the values of formant frequencies, and the area of the polygon represents the output acoustic space.

Key words: Vowel database, vowel polygons, acoustic space, MFCC Algorithm.

•**Associate Professor, Department of computer and automatic control, Faculty of mechanical and electrical engineering, Tishreen University, Lattakia, Syria.**
••**Postgraduate Student, Department of computer and automatic control, Faculty of mechanical and electrical engineering, Tishreen University, Lattakia, Syria.**

مقدمة:

يعتبر وضوح الكلام المنطوق بوساطة أي جهاز آلي للنطق مقياساً لكفاءته، وعلى الرغم من أن معظم الأجهزة المستخدمة حالياً لقراءة النصوص قد تنتج كلاماً مفهوماً، لكنه آلي وبعيد عن خصائص الكلام الطبيعي. لذا برزت الحاجة الملحة إلى توفر مولد آلي للنبرة عالي الجودة فينظم توليد الكلام، وهناك العديد من الأبحاث والمشاريع الحديثة في هذا المجال، ونظراً لتواضع الأعمال المُنجزة بالنسبة للغة العربية، قدم هذا البحث المراحل المتبعة لإنجاز قاعدة معلومات صوتية نبرية للهجات السورية كتلك المتوفرة للغات الأجنبية كقاعدة المعلومات (TIMIT) للإنكليزية و(BDSONS) للفرنسية[1].

وتشكل قواعد بيانات الكلام (Speech Database) الركن الرئيسي في بناء نظم حاسوبية مختلفة كنظم التعرف الآلي على الكلام والنطق الآلي والتعرف على المتحدث والتعرف على اللغات واللهجات، وتشكل البنية التحتية لبناء نظم تخاطب مع الحاسب، تسهم في توسيع دائرة مستخدمي الحاسب لتضم المكفوفين والأميين وذوي الاحتياجات الخاصة، إضافة لإمكانية التعامل مع الحاسب عبر شبكات الهواتف المختلفة، واستخدام تقنيات البصمة الصوتية كمفتاح للدخول إلى المواقع والبيانات الخاصة.

وتتكون قواعد بيانات الكلام عادة من ملفات صوتية (Wave Files) سبق أن سجلت لأشخاص باللغة المراد التعامل معها [2].

وكلما كانت قاعدة بيانات الكلام شاملة في محتواها وغنية بتنوع عناصرها، كلما ساعد ذلك على إخراج نظم حاسوبية ذات جودة عالية في أدائها.

أهمية البحث وأهدافه:

إنَّ أهم العقبات التي تواجه برامج التعرف الآلي على الكلام هي عدم الدراية بمساحات المجالات الصوتية الخاصة بكل لهجة؛ بسبب عدم توفّر قاعدة بيانات كافية للصوتيات وبالتالي عدم دراسة الخصائص الصوتية لكل لهجة.

يسمح هذا البحث بتطوير وسيلة مناسبة لفهم الكلام السوري بلهجاته العاميَة المتنوعة آلياً؛ وذلك من خلال معرفة مساحة المجال الصوتي السوري لكل لهجة والتي هي فعلياً مساحة المضلع الصوتي الممثل لمضلع هندسي رؤوسه هي قيم ترددات الحروف الصوتية السورية.

يهدف البحث إلى إنشاء قاعدة بيانات للصوتيات السورية، تشكل عنصراً أساسياً لبناء نظم حاسوبية مختلفة كنظم التعرف الآلي على الكلام والنطق الآلي والتعرف على المتحدث والتعرف على اللغات واللهجات، واستثمار قاعدة البيانات الصوتية المُنجزة في تطوير أدوات مساعدة لتحليل الكلام بالاعتماد على تحليل النغمة، عن طريق توليد وتحليل المضلعات الصوتية السورية.

طرائق البحث وموارده:

تضمن البحث توليد وتحليل مضلعات الحروف الصوتية السورية، أي أنه يتم تصنيف الحروف العربية إلى صامتة وإلى صوتية،وحساب معاملات السرعة والتسارع والانحراف المعياري للحروف الصوتية من أجل تشكيل مصفوفات الترابط، كما يتم حساب ترددات رنين المجرى الصوتي (Vocal Tract) والتي تحسب بالاعتماد على طيف

ترميز التنبؤ الخطي LPC(Linear Predictive Coding)، هذه الترددات موافقة لقمم الاستجابة الترددية لطيف المجرى الصوتي، وذلك لتقدير ترددات توافقيات المجرى الصوتي المعروفة بالنغمات (Formants)، وتستخدم الشبكات العصبونية ذات التغذية الأمامية للتدريب والاختبار للملفات الصوتية المشكلة لقاعدة البيانات الصوتية، حيث تشكل الحروف الصوتية المُسجلة في قاعدة البيانات الصوتية دخل للشبكة العصبونية، مع الأخذ بعين الاعتبار أن استخلاص السمات كان وفق خوارزمية معاملات تردد الميل MFCC(Mel Frequency Cepstral Coefficients).

وبذلك قدم البحث أدوات مساعدة لتحليل الصوتيات السورية، عن طريق توليد المضلعات الصوتية التي تمثل المجال الصوتي للهجات سورية، والتي تم إنشائها في بيئة تطويرية مناسبة وهي برنامج (Matlab) .

1-المتحدثون:

يتميز موقع الجمهورية العربية السورية باتساع مساحتها حيث تبلغ 185180 كيلو متر مربع، إضافةً إلى عدد السكان الذي وصل عام 2013 إلى حوالي22.5 مليون نسمة بمعدل نمو سكاني بلغ 0.15%.

ويقيم حوالي 56.1% من مجموع السكان في المدن، وتعتبر مدينة حلب أكبر المدن السورية، بينما دمشق الكبرى أكبر تجمع سكاني، ويقطن في المنطقتين 44% من مجموع الشعب إلى جانب المدن التي يفوق عدد سكانها المليونين وهما دمشق وحلب، فإن المدن التي يفوق عدد سكانها المليون أو شارفت على تخطي عتبته هي حمص، وحماه، واللاذقية، أما سائر المدن فهم يعيشون في المدن المتوسطة أو الصغيرة والريف إلى جانب وجود مجموعات قليلة من البدو في بادية الشام[3].

ومن ثمَّ فإن هذه العوامل: المساحة وعدد السكان تعني من الناحية اللغوية أن سكان سوريا يتسمون بالثراء والتنوع، فالجمهورية العربية السورية مكونة من أربعة عشر محافظة في سوريا، (دمشق، حلب، حمص، حماة، اللاذقية، طرطوس، درعا، السويداء، القنيطرة، إدلب، الحسكة، الطبقة، الرقة، ديرالزور) .

كل واحدة من هذه المحافظات لها لَهْجةٌ ولكْنَةٌ تختلف عن بقية المحافظات، بل إن كل مدينة ومنطقة تابعة لكل محافظة أيضاً لها، لَهْجةٌ ولكْنَةٌ وألفاظ تتميز بها عن بقية المناطق وعن المحافظة التابعة لها، بل إن هناك في بعض المدن والمحافظات في داخلها أحياء تختلف لهجاتها عن أحياء مجاورة لها، هذه ظاهرة موجودة في سوريا[4].

ويمكن تقسيم اللهجة السورية إلى[5]:

- اللهجات الغربية المصنفة ضمن اللهجة الشامية:
  - لهجة حلبية، تنتشر في مدينة حلب، وريفها نحو إدلب، هناك بعض التمايز اللفظي لريف محافظة إدلب.
  - لهجة جبلية، تنتشر في جبال محافظة اللاذقية، ومحافظة طرطوس، تتميز بلفظ القاف خلافًا لأغلبية الباقي.
  - لهجة ساحلية، تنتشر في مدن الساحل السوري : اللاذقية، وجبلة، وبانياس، وطرطوس؛ وتزداد تأثراً باللهجة اللبنانية كلما اتجهنا نحو الجنوب (طرطوس الأوفر تأثراً باللبنانية).
  - لهجة حمصية، تنتشر في حمص وريفها، تتميز بضم الحرف الأول من الكلام، وإبدال الألف أو التاء المربوطة ياء ساكنة في آخره مثل كبيري،لُعبي.
  - لهجة دمشقية، المنتشرة في مدينة دمشق وريفها، تتميز مع اللهجة الساحلية بمطّ الألف مثل لكان، تنتشر في حماه أيضاً.
  - لهجة أردنية، وهي المنتشرة في محافظة درعا، والجزء الأكبر من محافظة القنيطرة.

- لهجة حورانية، وهي اللهجة المنتشرة في محافظة السويداء، بشكل أساسي تتميز بلفظ الحروف العربية كالقاف والضاد.
  - اللهجات الشرقية المصنفة ضمن اللهجة العراقية:
- اللهجة العربية النجدية بين القبائل السورية، خصوصاً في دير الزور .
- لهجة شاوية، في محافظة الرقة، وشرق محافظة حلب في منبج وما جاورها.
- لهجة ماردلية، في محافظة الحسكة، وهي الأوفر محافظة على الآثار السريانية؛ وتعتبر جزءاً من عربية الجزيرة الفراتية في قسميها العراقي والسوري.

1-1- اختيار المتحدثين:

عند أخذ عينة لتمثيل مجموعة معينة فإنه من المهم أن يكون التمثيل دقيقاً إلى أقرب درجة ممكنة، وفي حالة هذا البحث فإنَّ أفضل تمثيل هو اختيار عينات لكل لهجات سوريا، وبما أنه لم يسبق إنجاز خارطة لغوية للجمهورية العربية السورية فإنه لا يمكن تحديد عدد اللهجات والمساحة الجغرافية التي تغطيها كل لهجة [6]، لذا كان الخيار البديل هو اختيار عينات من مناطق مختلفة يمثل سكانها لهجات مختلفة، بحيث تغطي اللهجات الرئيسية المتنوعة في سوريا؛ وكانت النتيجة القاعدة المعلوماتية التالية المعبرة عن عناصر قاعدة البيانات المنجزة:

الجدول (1) قاعدة بيانات معلوماتية عن متحدثي النظام.

الرقم المتسلسل	الاسم الثلاثي	سنة الميلاد	الجنس	الموطن
1	ابتسام أحمد إبراهيم	1993	أنثى	حلب/ مساكن هنانو
2	لافا عمر إسماعيل	1994	أنثى	دمشق/ المزة
3	هيوا صالح طاهر	1994	أنثى	القامشلي/ شارع السياحي
4	رحمه عبد الجبار سنون	1995	أنثى	إدلب/ معرة مصرين
5	شادي رجا مطر	1995	ذكر	السويداء/ القريا
6	نزار مطانيس ملحم	1992	ذكر	الحسكة/ القضاة
7	محمد أكرم الجباوي	1992	ذكر	درعا/ نوى
8	سامي هيثم أبو بالا	1990	ذكر	اللاذقية/ بسنادا
9	شذا ممدوح بيضون	1990	أنثى	حلب/ الأشرفية
10	رند توفيق حاج إبراهيم	1990	أنثى	اللاذقية/ جبلة
11	شادي غازي نعنوع	1986	ذكر	دمشق/ حرستا
12	ريم طاهر حيدر	1990	أنثى	طرطوس/ الغمقة

13	ريم جمال الدين الأخرس	1991	أنثى	حماه/ حي المدينة
14	غزوان نزيه زين الدين	1995	ذكر	السويداء/ قنوات
15	علي عزت كناج	1991	ذكر	طرطوس/ ميعار شاكر
16	محمد سلمان سعيد	1991	ذكر	اللاذقية/ القرداحة

الأشخاص يتحدثون لهجات هذه المدن كلهجة أم، جميعهم من السكان الأصليين لهذه المدن، ومن والدين ينتمون لهذه المدن وقضوا طفولتهم فيها، وجميعهم عاشوا أكثر من 16 عام في المدن المذكورة، جميع الأشخاص الذين يتم اعتماد تسجيلاتهم الصوتية هم من خلفيات ثقافية وحاصلين على درجات أكاديمية في مجالات مختلفة من جامعة تشرين، ولا يعانون من أي عيوب في النطق، وليس لديهم مشاكل سمعية ولا لفظية.

عدد الأشخاص المساهمين بتشكيل قاعدة البيانات هو 16 شخص (8 ذكور و 8 إناث)، أي أنه يتم دراسة أثر الجنس، أعمار هؤلاء الأشخاص تتراوح بين 20 و29 عام، (حيث تمَ اختيار الأشخاص من طلاب الجامعة)، أي سيتم تجاهل أثر العمر.

تمَ اختيار المشتركين بناءً على معايير جغرافية واجتماعية، حيث تمَ الأخذ بعين الاعتبار أخذ عينات من كافة المناطق السورية لإغناء قاعدة البيانات بلهجات متنوعة، ومن الجدير بالذكر أن الأشخاص الممثلين لكل لكنة يتحدثونها بشكل جليّ وبدون أن يكون هناك أي إشابة للهجاتهم بمفردات لهجات أخرى، وبالتالي من حيث التحديد الجغرافي شكل هؤلاء المشاركون عينات متجانسة من حيث طريقة النطق المناطقية.

1-2-التنسيق مع المتحدثين:

هناك عدة طرق لجمع عينات للكلام من المتحدثين منها[7]:

- أن يسافر الباحث إلى المتحدث في مقر إقامته مصطحباً معه أجهزة التسجيل ويقوم بالتنسيق مع المتحدث وتسجيل كلامه بعد إعداد الأجهزة.
- أن يسافر المتحدث من مقر إقامته إلى استديو خاص بالتسجيل.

ولكل من الطرق المذكورة محاسنه وعيوبه، ولكن الأنسب والأدق علمياً أن يتم دعوة المتحدثين إلى استديوهات معزولة صوتياً وإجراء عملية التسجيل فيها، وفعلاً تم إجراء العديد من التدابير للحصول على الموافقة بشأن دخول الإذاعة، وفي الموعد المحدد من قبل إدارة إذاعة أمواج في محافظة اللاذقية تمَّ اصطحاب المتحدثين إلى الاستديو والبدء بإجراءات التسجيل.

ونظراً لتواضع عدد عناصر قاعدة البيانات الصوتية المُنجزة بالمقارنة مع اللهجات المنتشرة على امتداد الجمهورية العربية السورية، وسعياً لجعل هذا البحث نواة لأبحاث متقدمة أكثر في هذا المجال، فقد تمَ إعادة كافة إجراءات التسجيل التي تمت في الإذاعة لكل متحدث ولكن بشروط تسجيل عادية، وباستخدام برنامج (Audacity)على الحاسب المحمول.

2-المادة اللغوية:

يعتمد مبدأ عمل نظم التعرف الآلي على الكلام على تدريب النظام باستخدام الموجات الصوتية اللغوية وما يقابلها من حروف أو رموز لغوية، ويتولى النظام الربط بينهما حتى يتعرف تلقائياً على الخصائص الصوتية لكل

صوت لغوي، وحيث أن الصوت اللغوي يظهر بعدة أشكال حسب موقعه في الكلمة والأصوات المجاورة لها وجنس المتكلم وعمره وغيرها، لذا لابد أن تتميز المادة اللغوية بالتنوع بحيث تغطي أكبر عدد من احتمالات ورود الصوت اللغوي[8,9].

2-1-المادة اللغوية المعتمدة:

تتنوع المادة اللغوية الملفوظة بين دراسة وأخرى، فأحياناً يتم لفظ الحرف الصوتي بشكل منفصل ولكن بذلك يُهمل تأثير الحرف المجاور للحرف الصوتي على هذا الصوتي،وأحياناً يتم لفظ الصوتي ضمن مقطع ونادراً أن يتم لفظ كلمة كاملة أثناء تسجيل قاعدة بيانات صوتية وذلك لأن الكلمة هي عبارة عن مجموعة مقاطع وبالتالي دراسة المقطع تفي بالغرض دون الحاجة لدراسة الكلمة كاملة، لذلك تمَّ في هذا البحث اعتماد المقاطع الصوتية كمادة لغوية.

ونظراً لأنَّ حوالي من 60% إلى 70% من الكلام السوري هو صوتي حيث تكثر الحروف الصوتية وحركات التشكيل النحوية في هذا الكلام، ونظراً لأن الحروف الصوتية تكون أكثر طاقة طيفية في التمثيل الهيستوغرامي(Histogram) وبالتالي تظهر بوضوح في الرسم الطيفي، وبالتالي تعتبر المقاطع الصوتية خير مُقيّم لخصائص اللهجات الملفوظة.

المقاطع الملفوظة المختارة هي مقطع مفتوح (ينتهي بحرف صوتي CV)، ومقطع مغلق (ينتهي بحرف صامت CVC)، وهناك أنواع أخرى من المقاطع في اللغة العربية ولكن تمَّ الاكتفاء بدراسة هذه المقاطع، حيث حرف V يرمز إلى الصوتي أو (Vowel) (تمَّ الأخذ بعين الاعتبار كل صوتيات العربية الطويلة "ا وي" أو " AAUUII" والصوتيات القصيرة "فتحة ضمة كسرة" أو " AUI")، أما حرف C يرمز إلى الصامت أو (Constant) وفي هذه الدراسةتمَّ اختيار حرف السين، والسبب في ذلك أن حرف السين له نهايات واضحة في طيف المقطع وبالتالي يسهل استخلاص الصوتي وفصله عن الصامت.

وبناءً على ما سبق سيقوم أشخاص قاعدة البيانات بلفظ المقاطع بالصيغة التالية:

سُوس، سَاس، سِيس

سُس، سَس، سِس

سُو، سَا، سِي

سُ، سَ، سِ

ونظراً لوفرة الدراسات المرجعية التي اعتمدت حرف الدال كحرف صامت في المقاطع الملفوظة لغرض تشكيل قاعدة بيانات، تم التسجيل كذلك للمقاطع التالية:

دُود، دَاد، دِيد

دُد، دَد، دِد

دُو، دَا، دِي

دُ، دَ، دِ

هذه المقاطع كُتبت على أوراق A4، بعد أن ينتهي الشخص من تسجيل المقاطع الأولى (سُوس، سَاس..) يسجل المقاطع الثانية (دُود، دَاد..) ثم يعيد عملية القراءة مرة ثانية وعندما ينتهي كل شخص من المهمة الموكلة إليه يتم حفظ ما قام به بملف بلاحقة (Wav)باسم هو الرقم التسلسلي لهذا الشخص والوارد في قاعدة البيانات المعلوماتية

لأشخاص النظام، ثم يُطلب من الشخص القيام بكل ما قام به مرة ثانية ولكن هذه المرة للتسجيل عبر برنامج (Auducity) عن طريق الحاسب المحمول.

كل شخص من الأشخاص الست عشر يلفظ المقاطع الأربع والعشرين جميعها مرتين (16*24*2= 768مقطع تسجيل ناتج) في الإذاعة ومثلها كذلك عن طريق الحاسب المحمول (عدد المقاطع الناتجة 1536)، يتم اللفظ بشروط عادية من السرعة ومستوى رفع الصوت.

3-التجهيزات:

3-1- شروط التجربة الأولى (التسجيل الاحترافي في استوديو معزول صوتياَ):

تمَ التسجيل داخل استديوهات معزولة صوتياً، حيث أن جدران الاستديو وأرضيته وسقفه معالجة بمواد ماصة للصوت كما تبطن الفراغات بمثل هذه المواد، كذلك الاستوديو خالياً من الشبابيك التي قد تسمح بدخول الضوضاء، وباب الاستوديو عازل للصوت أيضاً.

ويتصل الاستوديو بغرفة المراقبة من خلال العديد من الأجهزة، توجد في الحائط المتصل بغرفة المراقبة نافذة كبيرة من الزجاج، يراقب من خلالها الشخص الذي يشرف على التسجيل ما يقوم به الشخص الذي يقوم بالتسجيل، ويتم التسجيل لكل شخص على حدا، حيث يقرأ المقاطع المكتوبة على ورقة A4، بشروط عادية من السرعة وعلو الصوت. يتم الاستعانة بالأدوات التالية أثناء التسجيل في الاستديو:

- ميكروفون ديناميكي نوع Yamaha
- ميكسر Soundcraft Digital
- برنامج التسجيل Adobe Audition 3
- القناة الصوتية Stereo Channel
- تردد أخذ العينات 44100 Hz

جميع الأشخاص أنهوا التسجيل بجلسة واحدة، حيث استمر التسجيل لمدة ساعة واحدة.

3-2-شروط التجربة الثانية (التسجيل العادي في ظروف الوسط الطبيعية):

تمَ استخدام برنامج Audacity وذلك لكونه يعمل مع جميع صيغ ملفات الصوت، وله واجهة بسيطة سهلة الاستخدام، وله مميزات عديدة تساعد على تسجيل الأصوات بجوده عالية.

جميع الأشخاص أنهوا التسجيل بجلسة واحدة، حيث استمر التسجيل لمدة ساعة، حيث تمت عملية التسجيل هنا بالتزامن مع عملية التسجيل ضمن الاستديو المعزول، فكل شخص ينهي التسجيل في الاستديو وينتقل مباشرةً ليتم المهمة عن طريق الحاسب المحمول.

استغرق تسجيل كل شخص حوالي 50 ثانية، كذلك تردد أخذ العينات 44100 Hz.

4-الترميز الصوتي:

من الأعمال الأساسية في هذا البحث إجراء الترميز الصوتي للملفات الصوتية، ويتطلب الترميز الصوتي توفر القدرة على التركيز والمتابعة الدقيقة للأصوات اللغوية، وهي من الأعمال المرهقة ذهنياً لما تتطلبه من جهد وتركيز ، إذ يقوم بالاستماع للملف الصوتي أكثر من مرة ثم كتابة الرمز الصوتي المعبر عنه ثم الاستماع له ومطابقته بما كتب هو وتصحيح ما قد يحتاج للتصحيح.

وبناءً على ذلك تمَّ اعتماد أسماء للمقاطع الصوتية المسجلة (عملية الترميز الصوتي)، وهو عبارة عن رقم مكون من سبع خانات كما يلي:

الجدول (2) الترميز الصوتي.

الخانة التي تدل على مكان التسجيل	(1) إذا كان التسجيل ضمن استديو الإذاعة، (0)إذا كان التسجيل بواسطة الحاسب المحمول.
الخانة التي تدل على جنس المشترك	(1) إذا كان ذكر، (0) إذا كانت أنثى.
الخانة المعبرة عن نوع المقطع أي هل هو مفتوح CV أو مغلق CVC	(1) إذا كان CVC ، (0) إذا كان CV
الخانة المعبرة عن ماهية الصوتي المستخدم في المقطع	(1) إذا كان الواو ، (2) إذا كان الألف، (3) إذا كان الياء، (4) إذا كان الضمة، (5) إذا كان الفتحة، (6) إذا كان الكسرة.
الخانة المعبرة عن ماهية الصامت المستخدم في المقطع C	(1) إذا كان حرف السين ، (0) إذا كان حرف الدال.
الخانة المعبرة عن رقم التكرار	(1) في حال التكرار الأول، و(0) في حال التكرار الثاني.
الخانة المعبرة عن رمز اللهجة التي يمثلها المشترك	حلبية(1)، شامية(2)، حسكاوية(3)، إدلبية (4)، سويداوية(5) درعاوية (6)، لادقانية (7)، طرطوسية (8)، حموية (9)

وفي حال تطابق الخانات السبعة لمُتحدثَين، كما هو الحال بالنسبة للهجات حلب والسويداء واللاذقية، حيث تحوي قاعدة البيانات الصوتية على شخصين متطابقين بالجنس ويتحدثون نفس اللهجة، فإنه يتم إضافة الرمز a إلى يسار اسم الملف.

النتائج والمناقشة:

تضمن البحث عدة مراحل، فتمَ في المرحلة الأولى تحديد خصائص الحروف الصوتية وجمع الخوارزميات المناسبة لمعالجة الخصائص المقترحة، ثم بناء قاعدة بيانات لأصوات أشخاص سوريين تكون مناسبة لتصميم المشروع، ومن ثم بناء قاعدة بيانات تتضمن معلومات شخصية عن أشخاص النظام، حيث تمَ اختيار المشتركين بناءً على معايير جغرافية واجتماعية، فأُخذت العينات من كافة المناطق السورية لإغناء قاعدة البيانات بلهجات متنوعة، وتمَّ في المرحلة التالية إجراء تحليل MFCC على التسجيلات من أجل استخراج السمات، وبناء مصفوفات ارتباط من معاملات السرعة والتسارع والانحراف المعياري، ومن ثم بناء الشبكة العصبونية وتدريبها وحفظها، حيث اعتمدت الشبكة العصبونية كمصنف يقوم بإيجاد الرسم التخطيطي(Mapping) بين تسلسل أشعة السمة وبين عنصر الكلام المتعرف عليها، ثم إجراء الدراسة على نتائج تنفيذ الكود البرمجي للتسجيلات، حيث يتم رسم مضلعات الحروف الصوتية

السورية، وهي عبارة عن مضلعات هندسية رؤوسها تمثل قيم الترددات للحروف الصوتية السورية، وحساب مساحة هذا المضلع التي تمثل المجال الصوتي الخاص بكل لهجة سورية.

1- توليد المضلع الصوتي:

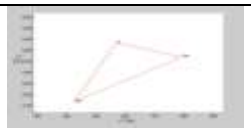
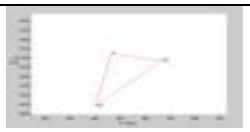
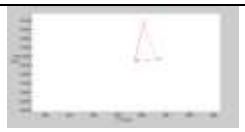
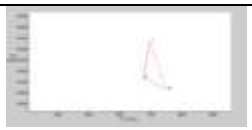
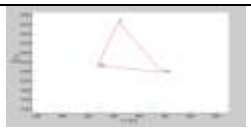
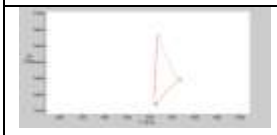
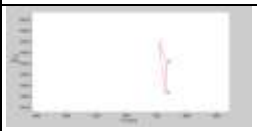
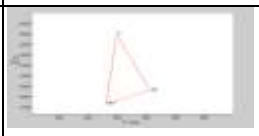
المضلع الصوتي هو شكل هندسي، يلزم لرسمه معرفة إحداثيات رؤوسه، وبما أن رؤوس المضلع الصوتي هي قيم ترددات نغمات الحروف الصوتية، فالعمل البرمجي ينصب على معرفة قيم ترددات هذه النغمات، وبما أنَّ النغمات هي إحدى سمات الكلام (Speech Features) فيتم استخلاصها وفق الخوارزمية MFCC وهي خوارزمية جاهزة برمجياً في (MathWorks) يتم تغيير بارامتراتها بما يتناسب مع عدد العينات وتردد التقطيع المستخدم في هذا البحث.

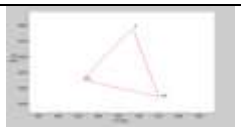
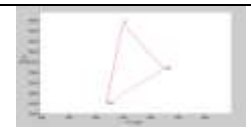
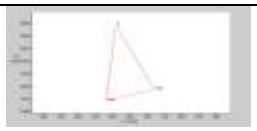
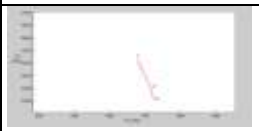
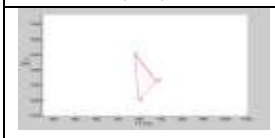
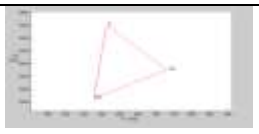
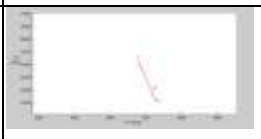
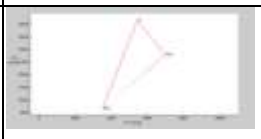
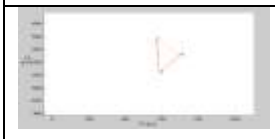
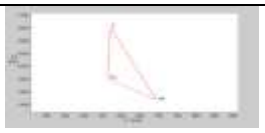
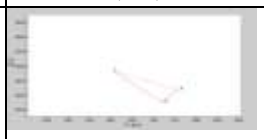
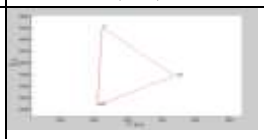
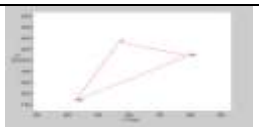
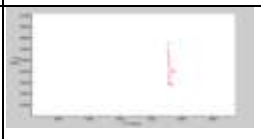
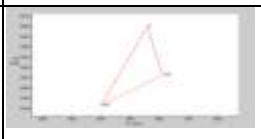
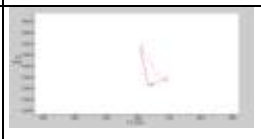
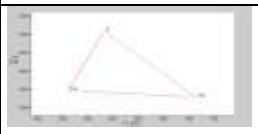
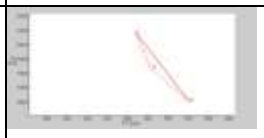
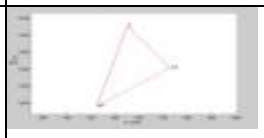
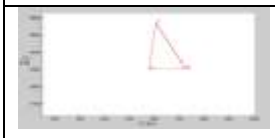
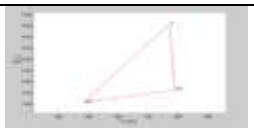
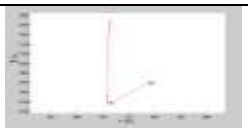
بنتيجة تطبيق خوارزمية MFCC نحصل على النغمات متضمنة العديد من البارامترات (Formant Parameters)مثل عرض الحزمة (Bandwidth) والميل (Slope) والتردد (Frequency)، ولحساب التردد تمَّ اعتماد خوارزمية LPC المتوفرة أيضاً في (MathWorks) بعد ضبط رتبة التنبؤ فيها بما يتناسب مع خصوصية العينات المأخوذة، وبنتيجة هذه الخوارزمية نحصل على قيمتي ترددي النغمتين F2،F1 لكافة الحروف الصوتية لجميع المتحدثين.

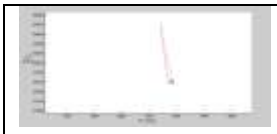
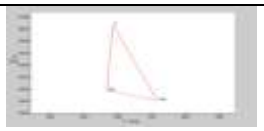
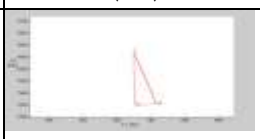
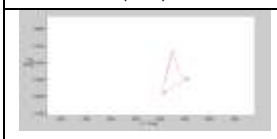
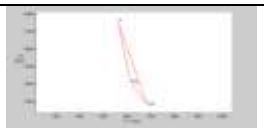
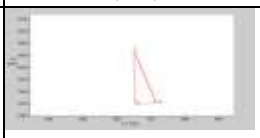
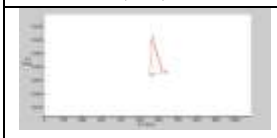
وبالتالي لرسم المضلع الصوتي الطويل الاحترافي (أي المضلع الخاص بالحروف الصوتية الطويلة وهي الألف والواو والياء والمرمزة بالحروف " AAUUII"، والذي تمَّ تسجيله في ظروف احترافية ضمن أستوديو معزول صوتياً) يتم استخدام التعليمة البرمجية (Plot) لرسم هذه النقاط الثلاثة في المستوي F2-F1، حيث أنه كل من النقاط "AAUUII " تُحدد بإحداثيين هما قيمتي F1 وF2، وبعد رسم رؤوس المضلع يتم الاستعانة بالتعليمة (Line) للوصل بخط بين هذه النقاط وتشكيل أضلاع المضلع، وبالتالي ينتج المضلع الصوتي، ولحساب مساحته تُستخدم التعليمة البرمجية (Polyarea)، حيث ما يهم البحث هو الحصول على قيمة هذه المساحة لأنها تمثل مساحة المجال الصوتي بغض النظر عن شكل المضلع الصوتي الناتج.

2-النتائج العملية:

وقد أعطى تطبيق الكود البرمجي النتائج التالية عند حساب مساحات المجالات الصوتية للحروف الصوتية الطويلة والقصيرة، وذلك في ظروف التسجيل الاحترافي داخل ستوديو معزول صوتياً وظروف التسجيل العادية ضمن شروط الوسط المحيط، والأشكال التالية توضح هذه النتائج:

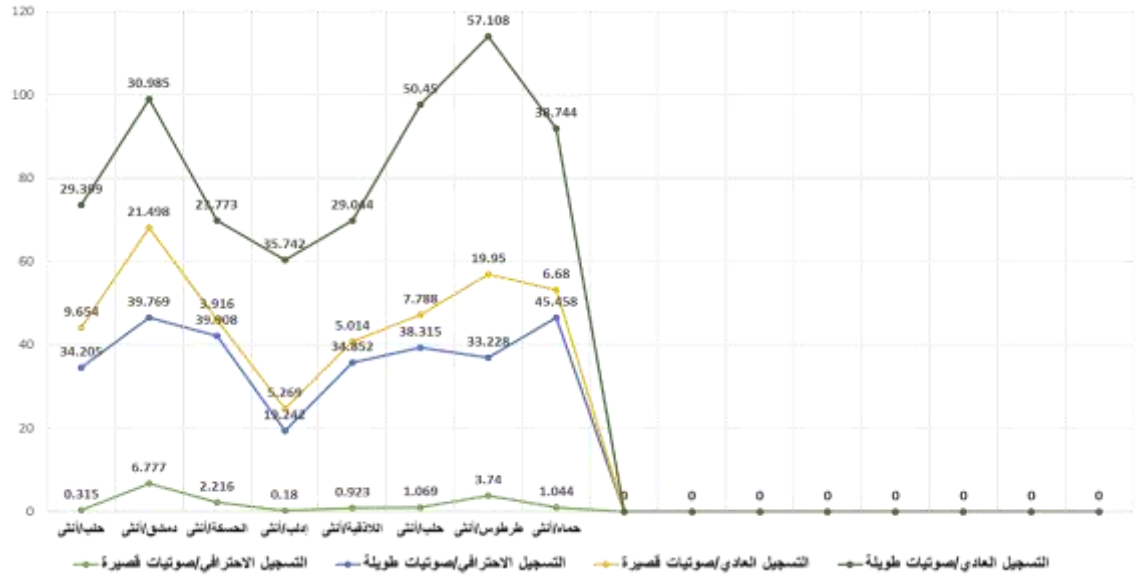
أنثى من حلب	 (1)	(17)	 (33)	 (49)
أنثى من دمشق	(2)	 (18)	 (34)	 (50)
أنثى من الحسكة	(3)	 (19)	 (35)	(51)

أنثى من إدلب	 (4)	(20)	 (36)	(52)
ذكر من السويداء	 (5)	 (21)	(37)	(53)
ذكر من الحسكة	(6)	(22)	(38)	 (54)
ذكر من درعا	 (7)	 (23)	 (39)	 (55)
ذكر من اللاذقية	 (8)	 (24)	 (40)	(56)
ذكر من السويداء	 (9)	 (25)	 (41)	(57)
ذكر من الحسكة	(10)	(26)	(42)	(58)
ذكر من درعا	(11)	 (27)	 (43)	(59)
ذكر من اللاذقية	(12)	 (28)	 (44)	 (60)
أنثى من حماه	 (13)	(29)	 (45)	(61)

ذكر من السويداء	(14)	(30)	(46)	 (62)
ذكر من طرطوس	 (15)	 (31)	(47)	 (63)
ذكر من اللاذقية	 (16)	 (32)	(48)	 (64)
	(أ)	(ب)	(ج)	(د)

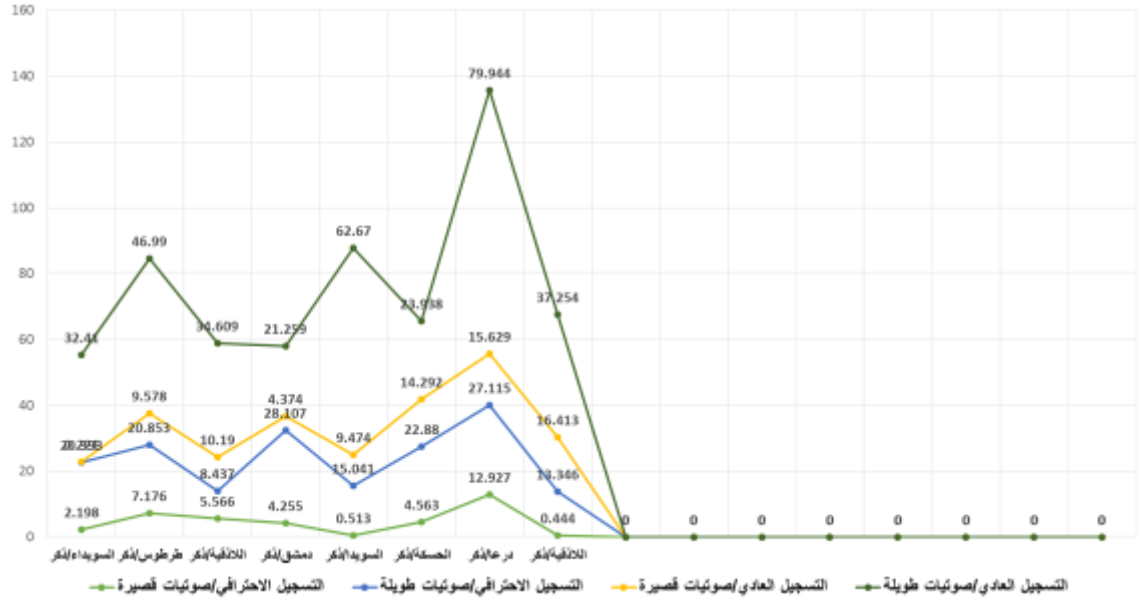
الشكل (1-64) المضلعات الصوتية الناتجة: (أ) المضلعات الصوتية الطويلة الاحترافية، (ب) المضلعات الصوتية القصيرة الاحترافية، (ج) المضلعات الصوتية الطويلة العادية، (د) المضلعات الصوتية القصيرة العادية

وبمقارنة هذه النتائج لإناث قاعدة البيانات، تبيّن أن مساحات المضلعات الصوتية قد سجلّت أعلى القيم بالنسبة إلى لهجات طرطوس ودمشق، وأدنى القيم بالنسبة إلى لهجة حلب، حيث أن الفرق الرئيسي بين هذه اللهجات هو في لفظ الأحرف الصوتية، حيث تُمال الألف إلى ياء في اللهجة الحلبية، ففي لهجة حلب تكثر جداً الألف المرققة، أما في لهجة الساحل تكثر الألف المفخمة تفخيماً تاماً، وأما في دمشق فتسود الألف المغلظة، أي أن اللهجة الدمشقية تقارب الساحلية من حيث التفخيم، ويظهر الشكل (65) المخطط البياني الممثل لمساحات المضلعات الصوتية لإناث قاعدة البيانات الصوتية المُنجزة، حيث تمَّ رسم جميع قيم المساحات الناتجة بمخططات (Excel) يعبر محورها الشاقولي عن قيمة المساحة الناتجة:



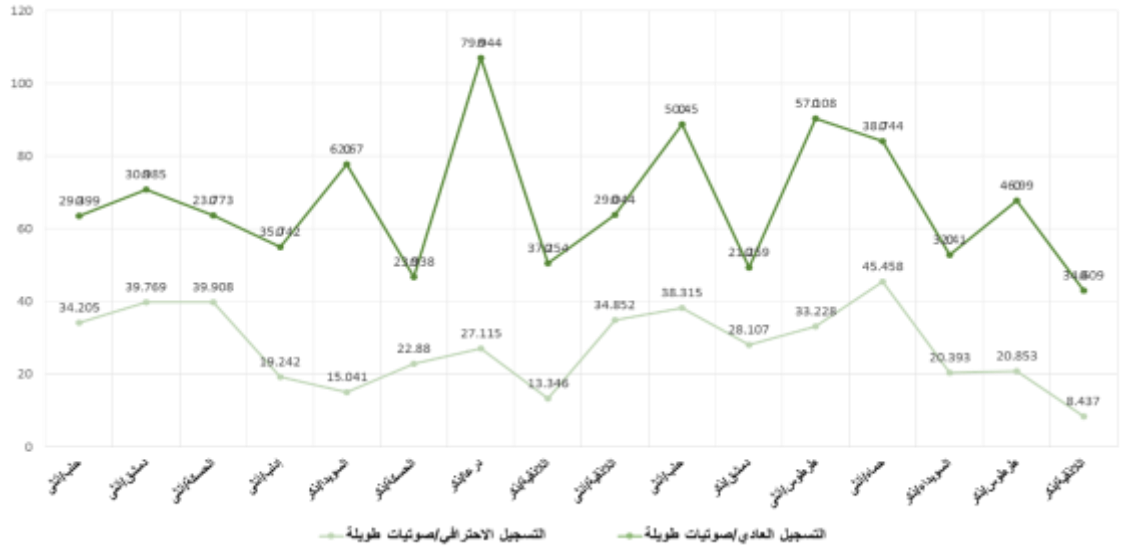
الشكل(65): مساحات المضلعات الصوتية لإناث قاعدة البيانات الصوتية المُنجزة

وبالنسبة لعناصر قاعدة البيانات الذكور تبيّن أن مساحات المضلعات الصوتية الطويلة والقصيرة في حالتي التسجيل الاحترافي والتسجيل العادي، قد بلغت أعلى القيم في حال لهجات درعا وطرطوس، كما يظهر في الشكل (66):



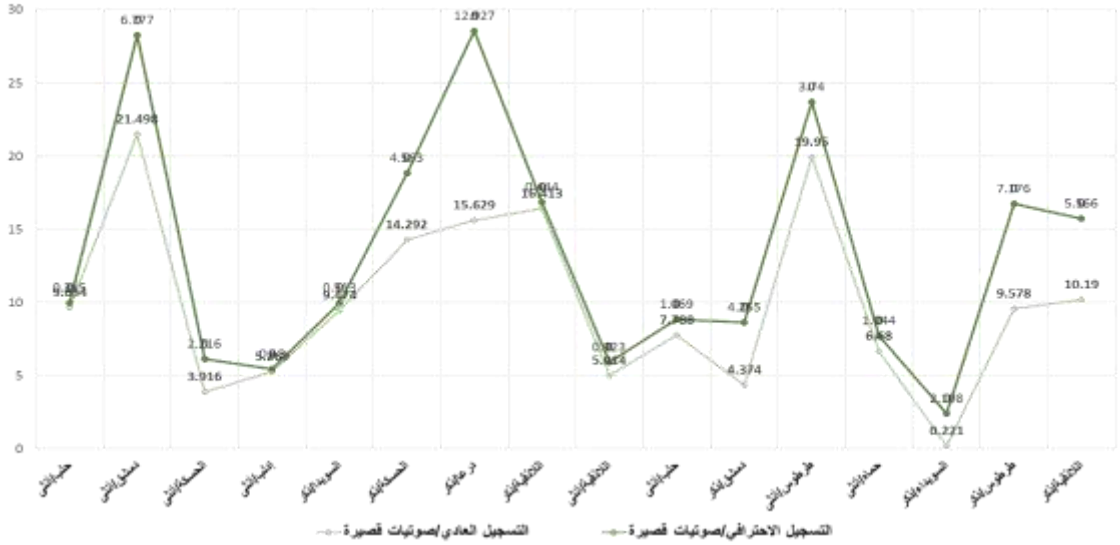
الشكل(66): مساحات المضلعات الصوتية لذكور قاعدة البيانات الصوتية المُنجزة

تمَّت مقارنة مساحات المضلعات الممثلة للصوتيات الطويلة الناتجة عن التسجيل العادي مع مساحات المضلعات الممثلة للصوتيات الطويلة الناتجة عن التسجيل الاحترافي، فأظهرت المقارنة أنَّ مساحات المضلعات الناتجة عن التسجيل العادي أكبر من مساحات المضلعات الناتجة عن التسجيل الاحترافي، ويُعلل ذلك بأنه أثناء التسجيل العادي أي التسجيل في ظروف الوسط الطبيعي تتداخل إشارة الضجيج مع إشارة الصوت، وبالتالي تكون قيم الترددات المُستخلصة من الإشارة المشوبة ذات قيم متباينة، تتوضع هذه القيم على مسافات متباعدة في فضاء المجال الصوتي، وبالتالي تشكل بمواضعها المتباعدة رؤوس مضلع ذو مساحة أكبر من المضلع الذي تشكل رؤوسه قيم منتظمة متقاربة فيما بينها، كما يظهر الشكل (67):



الشكل(67): تباين مساحات المضلعات الصوتية الطويلة في حالتي التسجيل العادي والاحترافي

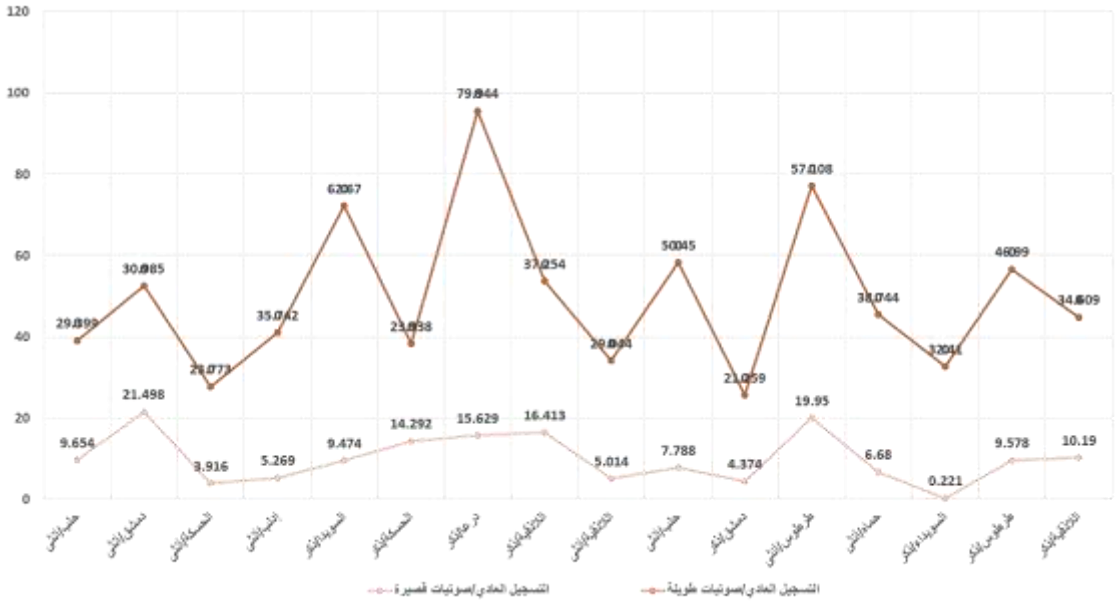
وقد تمَّ الوصول إلى نفس النتيجة عند مقارنة مساحات المضلعات الممثلة للصوتيات القصيرة الناتجة عن التسجيل العادي مع مساحات المضلعات الممثلة للصوتيات القصيرة الناتجة عن التسجيل الاحترافي، حيث أن مساحات المضلعات الناتجة عن التسجيل العادي أكبر من مساحات المضلعات الناتجة عن التسجيل الاحترافي، كما في الشكل التالي:



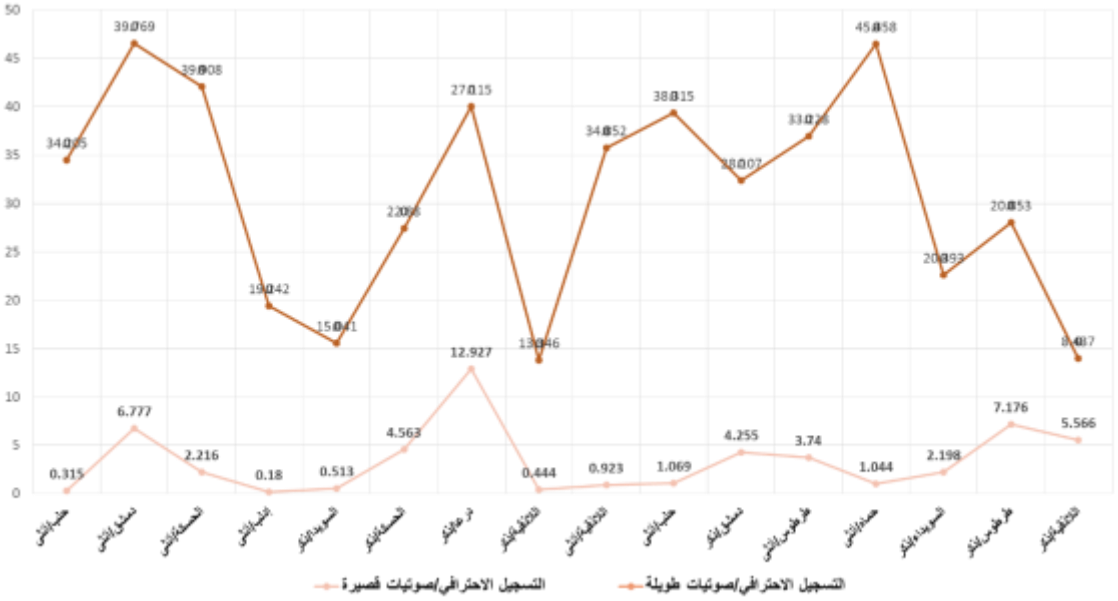
الشكل(67): تباين مساحات المضلعات الصوتية القصيرة في حالتي التسجيل العادي والاحترافي

إنَّ مساحة المجال الصوتي للصوتيات الطويلة أكبر من مساحة المجال الصوتي للصوتيات القصيرة، ففي حال التسجيل الاحترافي كانت مساحات مضلعات الصوتيات الطويلة أكبر بثمانية أضعاف من مساحات مضلعات الصوتيات القصيرة، في حين أنها كانت أكبر بأربعة أضعاف أثناء ظروف التسجيل العادية؛ ويُعزى اتساع المجال الصوتي للصوتيات الطويلة مقارنةً بمجال الصوتيات القصيرة إلى الاختلاف الواضح في القيم الترددية لنغمات الحروف

الصوتية الطويلة مما يؤدي إلى توليد مضلع بمساحات أكبر من المضلع الناتج عن الصوتيات القصيرة ذات القيم الترددية المتقاربة، كما تبين الأشكال (68) و (69):



الشكل(68): تباين مساحات المضلعات الصوتية الطويلة والقصيرة في حالة التسجيل العادي

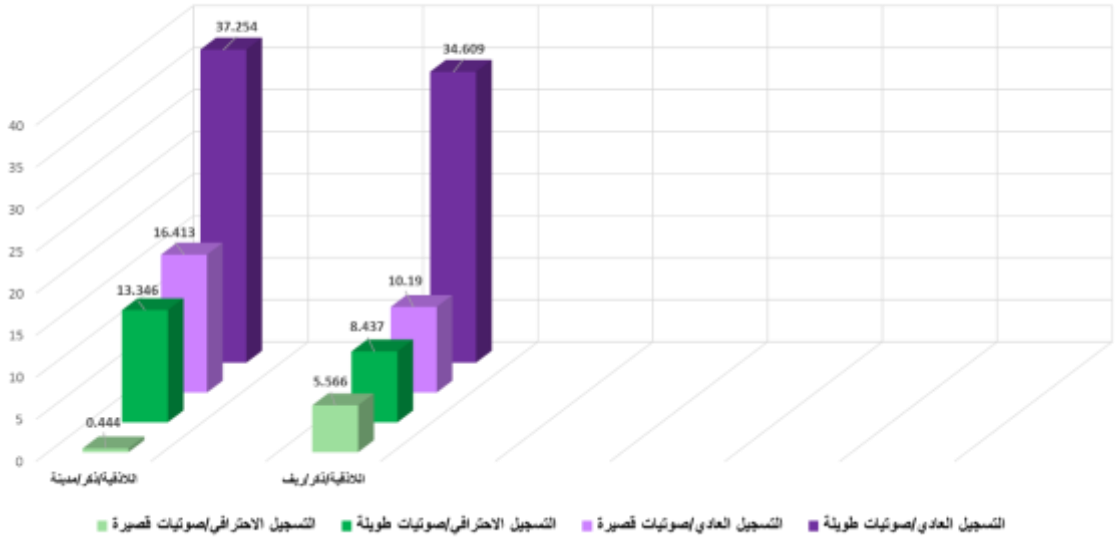


الشكل(69): تباين مساحات المضلعات الصوتية الطويلة والقصيرة في حالة التسجيل الاحترافي

تمَّ تسجيل تقارب كبير يصل لحدود التطابق بين مساحات المضلعات الصوتية الناتجة عن شخصين ذكرين يتحدثان نفس اللهجة، وهي لهجة محافظة اللاذقية، أحدهما من السكان الأصليين لريف اللاذقية والآخر من السكان الأصليين لمدينة اللاذقية.

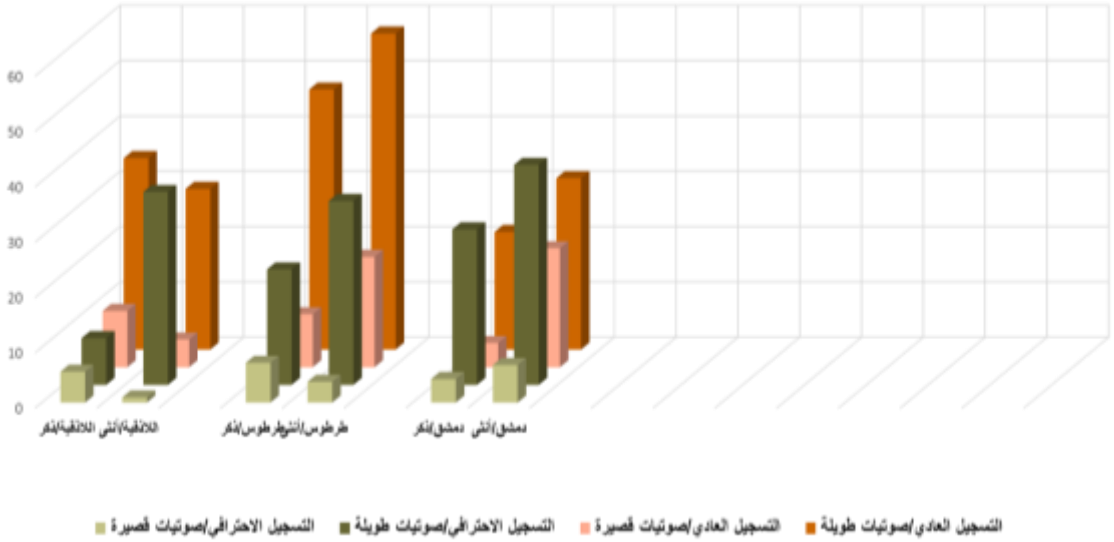
أي أنه بمقارنة عنصرين من قاعد البيانات الصوتية المُنجزة متطابقان بالجنس واللهجة، تبيّن أن المضلعات الناتجة عن التسجيل العادي متقاربة في المساحة لكلا العنصرين، وأن المضلعات الناتجة عن التسجيل الاحترافي

متقاربة في حال الصوتيات الطويلة، وتكون أكبر مساحةً بقليل للشخص الريفي من الشخص المدني في حال الصوتيات القصيرة، ويمكن تفسير ذلك بأن أصوات سكان الريف تكون أكثر حدية من أصوات سكان المدن، وبالتالي تكون قيم ترددات نغمات أصواتهم أعلى من مثيلاتها عند سكان المدن، كما يلي في الشكل (70):



الشكل(70): مقارنة مساحات المضلعات الصوتية لشخص من ريف اللاذقية مع شخص من مدينة اللاذقية

وتمّت دراسة وتحليل المضلعات الصوتية للصوتيات الطويلة والقصيرة والناتجة عن تسجيل احترافي وعن تسجيل عادي لعدة لهجات، بحيث يتحدث كل لهجة ذكر وأنثى، وبمقارنة النتائج تبيّن أنه في لهجات دمشق وطرطوس واللاذقية تكون مساحات المضلعات الصوتية الناتجة عن تسجيلات الإناث أكبر مساحةً من تلك المضلعات الناتجة عن تسجيلات الذكور، ويُبرر ذلك أن ترددات صوت الأنثى تكون أعلى من ترددات صوت الرجل لأن صوتها أرفع ويرجع ذلك إلى صغر الحبال الصوتية للأنثى، والشكل (71) يوضح نتائج هذه المقارنة:



الشكل(71): مقارنة مساحات المضلعات الصوتية لإناث وذكور اللهجة الواحدة

الاستنتاجات والتوصيات:

تضمنت هذه الدراسة إنشاء قاعدة بيانات للصوتيات السورية، تضمنت ست عشرة عنصراً من مناطق مختلفة من الجمهورية العربية السورية، كانت تلك العناصر متنوعة من حيث الجنس، فاحتوت ذكور وإناث من شرائح عمرية متقاربة، أي أنه تمَّت دراسة أثر الجنس على مساحة المجال الصوتي، وإغفال أثر العمر؛ لذلك يوصى بتوسيع قاعدة البيانات الصوتية المُنجزة بإضافة تسجيلات صوتية لأشخاص من فئات عمرية مختلفة، ودراسة تأثير عامل العمر على مساحة المجال الصوتي.

وتجدر الإشارة أنه سيكون من المفيد جداً بناء قاعدة بيانات الأصوات الهاتفية للمتحدثين السوريين، بحيث تمثل كافة شرائح المتحدثين السوريين من حيث اللهجة والجنس والعمر ووسائل الاتصالات المتوفرة في الجمهورية العربية السورية، وتكون الركيزة الأساسية لنظم التعرف على الكلام حيث تستخدم في تدريب محركاتها.

المراجع

[1] STANEK, M., SIGMUND, M. Speaker Dependent Changes in Formants Based on Normalization of Vowel Triangle. *In Proc. 23rd International Conference RADIOELEKTRONIKA. Pardubice*. Czech Republic, 2013, pp. 337-341.

[2] ALGHAMDI, M. *Analysis, Synthesis and Perception of Voicing in Arabic*. Al-ToubahBookshop, Riyadh. 2004, P. 50.

[3] KENSTOWICZ, M. *Parametric variation and accent in the Arabic dialects*, 1983, CLS19: 205-213.

[4] IRSHIED, O; MICHAEL, K..*Some phonological rules of Bani-Hassan Arabic: A Bedouin dialect*, Studies in the Linguistic Sciences 14,1984, 109-147.

[5] SIGMUND, M., ZELINKA, P., Analysis of Voiced Speech Excitation Due to Alcohol Intoxication. *Information Technology and Control*, 2011, vol. 40, pp. 145-150.

[6] STANEK, M., POLAK, L. Algorithms for Vowel Recognition in Fluent Speech Based on Formant Positions. *In Proc. 36th International Conference on Telecommunication and Signal Processing*. Rome (Italy), 2013, pp. 521-525.

[7] HU, R. and DAMPER,R. *Fusion of two classifiers for speaker identification: removing and not removing silence*, the 8th International conference on Information fusion,vol 1,2006, 429-436.

[8] ABDO, M. S., KANDIL, A. H., El-BIALY, A. M. Automatic detection for some common pronunciation mistakes applied to chosen Quran sounds. *Biomedical Engineering Conference (CIBEC), 5th Cairo International*, 2010, pp. 219-222.

[9] IQBAL, H. *On Vowels Segmentation and Identification Using Formant Transitions in Continuous Recitation of Quranic Arabic*, Chapter of "New Challenges in Applied Intelligence Technologies", Springer Berlin, 2008, pp. 155-162.