# نظام ترجمة لفظية للنصوص المكتوبة باللغة الانكليزية

**سهام حسن خيون**
**قسم علوم الحاسبات/كلية العلوم/ جامعة البصرة**

**الخلاصة Abstract:**

تعد عملية مراجعة البحوث( المكتوبة باللغة الانكليزية) والاطلاع عليها مهمة شاقة أمام العديد من الباحثين العرب من حيث الترجمة والقراءة هناك برامج عديدة عالجت مسألة الترجمة السريعة من الانكليزية إلى العربية لكن لا توجد بحوث أو دراسات أو برامج تساعد في عملية قراءة البحوث والكتب الانكليزية لتمكن الباحث من اللغة أولا ومن اللفظ السليم للحروف ثانيا وأخيرا من فهم واستيعاب ما جاء في النصوص.

لذلك ارتأينا تصميم برنامج له القدرة على قراءة النصوص الانكليزية سواء كانت كتب أو بحوث أو ما شابه (التي تكون بشكل ملفات نصية Text File ) ومن ثم تكوين ملف صوتي(بصيغة Wav) وملف نصي (بصيغة Txt) يتضمن ترجمة لفظية للنص الانكليزي باللغة العربية.

تتلخص طريقة عمل البرنامج المقترح بإدخال اسم الملف النصي ومساره الخزني من قبل المستفيد ومن ثم يقوم البرنامج بتحليل مفردات النص المدخل بتوليد قائمة بكل الكلمات بعد هذه المرحلة يسلك البرنامج مسارين مختلفين للحصول على ملف صوتي من المسار الأول وملف نصي باللغة العربية من الثاني.
**المسار الأول:** وهنا يعمل برنامج فرعي يدعى المترجم اللفظي الصوتي يستعمل آلية بحث كفوءة للبحث عن المقطع الصوتي لكل مفردة في النص (بالاستعانة بأحد المعاجم الناطقة) وخزن مقطعها الصوتي في الملف الصوتي الذي يكونه البرنامج بشكل تراكمي عن طريق دمج المقطع الصوتي للكلمة الحالية مع بقية الملف وهكذا إلى نهاية قائمة الكلمات.

**المسار الثاني:** هنا يتم معالجة النص وفقا لمفهوم معالجة اللغات الطبيعية حاسوبيا حيث يعمل برنامج فرعي آخر يدعى المترجم اللفظي النصي على ترجمة النص الانكليزي إلى اللغة العربية لفظيا مولدا ملف نصي لفظي لكل النص المدخل يستند هذا البرنامج على معجم صمم لهذه المرحلة يحتوي جدول بكل الحروف والأصوات الانكليزية وما يقابلها بالعربية.

## 1. المقدمة Introduction

إن الدراسات والبحوث والكتب تتزايد يوما بعد آخر في كل المجالات العلمية والأدبية على حد سواء ويعاني الكثير من الباحثين من مواكبة هذا التطور من خلال مراجعة البحوث والدراسات والكتب ذات الصلة بموضوع بحثه واحد أهم الأسباب هو عدم تمكن العديد من الباحثين من اللغة الانكليزية قراءة وكتابة ومعاني نتيجة طرق تدريس اللغة الانكليزية في المدارس والجامعات العراقية التي تكاد تكون غير مجدية.

هناك العديد من البحوث والدراسات والبرامج التي تساعد في ترجمة النصوص الانكليزية إلى اللغة العربية لكن لا توجد دراسات وبحوث تساعد في عملية قراءة النصوص الانكليزية ليتمكن الباحث من نطق الأصوات بصورة سليمة. هناك العديد من القواميس الناطقة تساعد في عملية القراءة لكنها تتطلب جهد ووقت كبير من الباحث لأن عملية الترجمة اللفظية تتم بإدخال كلمة – كلمة فما بالك بكتاب أو بحث وكمثال لهذا النوع قاموس Babylon، توجد بعض القواميس الناطقة لها القدرة على قراءة جملة مرة واحدة مثل برنامج TALK IT! لكن تبقى مسألة الوقت والجهد عائق أمام كثرة الاستخدام، لذلك ارتأينا تصميم نظام له القدرة على قراءة النصوص الانكليزية بيسر وسهولة مهما كان حجم النص كبير كما أضفنا مسألة الملف النصي الناتج من الترجمة اللفظية لزيادة توكيد دقة لفظ الأصوات الانكليزية.

## 2. معالجة اللغات الطبيعية حاسوبيا Natural Language Processing

كما لا يخفى على الكثير إن معالجة اللغات الطبيعية واحدة من أهم فروع الذكاء الاصطناعي والهدف منها جعل الحاسوب يفهم اللغة الطبيعية كتابة وقراءة حتى نتمكن من التعامل مع الحاسوب بلغتنا الطبيعية بدلا من لغات البرمجة لأن اللغة الطبيعية وفهمها احد مميزات الإنسان عن المخلوقات الأخرى.

يعرف نظام معالجة اللغة الطبيعية بأنه دالة طبيعية متعددة القيم ذات اتجاهين بين مجموعة المعاني الطبيعية Meanings ومجموعة التراكيب اللغوية للمعنى الواحد( Edward Gibson, 1998).

يتكون نظام معالجة اللغة الطبيعية بصورة عامة من خمس منظومات تتفاعل فيما بينها بغض النظر عن اللغة الطبيعية التي يتم معالجتها، الشكل(1) يوضح مكونات منظومة معالجة اللغة الطبيعية. في بحوثنا السابقة كانت اللغة العربية الهدف أما في بحثنا هذا نقوم بمعالجة اللغة الانكليزية وبما إننا نعالج بحوث وكتب ونصوص موثقة وقد تم التأكد من صحة طباعتها من قبل مؤلفيها سوف نتعامل مع بعض من هذه المنظومات التي نحتاجها لحل المسألة المقترحة.

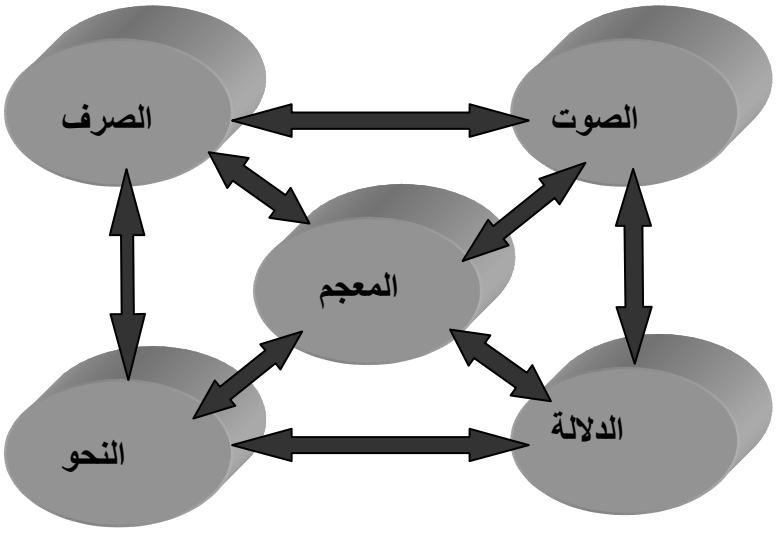
**1-2 الصوت Phonology :** الصوت اللغوي هو أكثر الظواهر تأصلاً .إن ظاهرة الكلام عبارة عن نشاط حي ديناميكي يتصف بالتدفق والتلون واللامحدودية ويتفاعل فيه، ومعه وبشكل حاد عناصر متعددة مثل الظرف والبيئة والخلفية والمتحدث والمستمع (نبيل علي، 1988). نظام معالجة الصوت مسؤول عن اخذ الأصوات وترجمة الإدخال إلى كلمات (Eugene Charniak, 1988).

**2-2 الصرف Morphology :** الصرف هو فرع اللسانيات الذي يتعامل مع البنية الداخلية للكلمات من حيث تكوين عناصرها الأولية والتفاعل بين هذه العناصر (نبيل علي، 1988)، إذ يهتم الصرف بالدرجة الأساس بمعالجة الكلمة بغض النظر عن موقعها في الجملة وهي أول مرحلة في معالجة اللغة الطبيعية حاسوبياً.

**3-2 النحو Syntax :** النحو عند اللسانين هو علم بأصول يعرف به تغيير أواخر الكلمات إعرابا وبناءً وغاية هذا العلم صون اللسان عن اللحن في الكلام وبه يستعان على فهم الكتاب (القرآن الكريم) والسنة النبوية وكلام العرب الأوائل (احمد قصير العاملي ، 1985)، أما عند علماء اللسانيات الحاسوبية فيعرف على انه دراسة بنية الجملة وذلك من حيث ترتيب عناصرها أو مكوناتها أو العلاقات التركيبية (البنائية) والوظيفية التي تربط بين هذه العناصر(نبيل علي، 1988).

**4-2 المعجم Lexicon :** هو عبارة عن قائمة بالكلمات التي ترتبط فيها كل كلمة من هذه الكلمات مع سماتها الدلالية والنحوية والصرفية فضلاً عن سمات أخرى (Gerald Gazdar, 1989). أصبحت أهمية المعجم تزداد يوماً بعد آخر ابتداءً من عقد الثمانينيات من القرن الماضي بسبب وجوده كصفة أساسية في كل أنظمة معالجة اللغات الطبيعية(سهام حسن، 2007).

**5-2 الدلالة Semantic :**المقصود بالدلالة تحديد معنى الجملة ،وهو العلم الذي يدرس الشروط الواجب توافرها في العنصر (اسم أو فعل أو حرف)حتى يكون قادراً على حمل المعنى، والتحليل الدلالي فرعان، الأول يهتم ببيان معاني الكلمات وهنا يرتبط مع المعجم، أما الثاني فيهتم بمعاني الجمل والعبارات. إن فهم معنى الجملة اللغوية يتشكل من تجميع معاني الكلمات المكونة للجملة (Edward Gibson, 1998). وان معنى الكلمة يعتمد على المعرفة المرتبطة بهذه الكلمة وايضاً بقية المعاني التي سترتبط بالكلمة أو بالجملة نفسها خلال المعالجة. وللدلالة دور مهم في رفع الغموض عن الكلمات أو الجمل(سهام حسن، 2007).




الشكل (1) منظومة معالجة اللغات الطبيعية

3. **اللغة الانكليزية English Language:**

تعتبر اللغة الانكليزية إحدى أهم اللغات الطبيعية في العالم لأنها تعد جواز سفرك لمعظم دول العالم فهي اللغة الثانية إن لم تكن الأولى، رغم ذلك فهي لغة حديثة ومتجددة دائما، تخضع لمقاييس لجنة دولية خاصة أكثر من سياقات اللغة نفسها.

اللغة الانكليزية ذات صفات تختلف عن اللغة العربية في نقاط وتشترك معها في أخرى نسبة لمنظومة معالجة اللغات الطبيعية آنفة الذكر. فهي لا تتمتع بخاصية الاشتقاق الصرفي مثل اللغة العربية مثلا فعل الكتابة ومشتقاته فيها (write) في حين كتاب (book) ومكتبة (library) أما بالنسبة للغة العربية فكل الكلمات ذات الصلة تكون مشتقة من جذر واحد، لو نأخذ المثال السابق تكون كلها مشتقة من الجذر (كتب) (كتب، كاتب، مكتوب، كتاب، مكتبة،...) لذلك يصعب معالجة اللغة الانكليزية من الناحية الصرفية. أما بالنسبة لقواعد لفظ الأصوات في اللغة الانكليزية هناك مجموعة من القواعد يمكن أتباعها والاستفادة منها في معالجة اللغة إلا انه توجد الكثير من الشواذ باللفظ عن القواعد الموضوعة ،أحيانا نجد إن الحرف يكتب ولا يلفظ ،وأحيانا أخرى يلفظ لكن بصوت آخر يعتمد على أمور مختلفة، لكن حروف العلة(i, o, e, a, u) في معظم الكلام تكون غير خاضعة لقواعد لفظية عامة لذلك كانت الجزء الأصعب في معالجة المسألة المقترحة. والجدول(1) يبين الأصوات الانكليزية وأنواعها المصنفة من قبل أهل اللغة (Michael Swan, 1995) وقمنا بإضافة رقم لكل نوع لتسهل عملية المعالجة.

الجدول (1) يوضح الأصوات الانكليزية

| English Sounds | | | |
|---|---|---|---|
| Vowels | | Consonants | |
| (4)Pure | (3)Diphthongs | (2)Voiced | (1)Voiceless |
| i | ei | b | P |
| i: | eə | d | T |
| e | ai | g | K |
| a | ou | ð | Θ |
| a: | au | v | F |
| o | iə | z | S |
| o: | oi | dʒ | Tδ |
| u | uə | ʒ | δ |
| u: | | m | H |
| ʌ | | n | |
| ə | | l | |
| ə: | | r | |
| | | ŋ | |
| | | w | |
| | | j | |

4. **الترجمة اللفظية Pronunciation Translation**

إن ظاهرة الترجمة كانت نتيجة الظاهرة اللغوية لدى البشر ، فما إن تفرق البشر إلى شعوب وقبائل وتطورت لديهم الظاهرة اللغوية حتى برزت الحاجة للترجمة لتوصيل الأفكار والمشاعر والرغبات ولتحقق التفاهم الذي هو الوظيفة العليا للغة(سلمان داود الوسطي، 2002).

إن الترجمة تأخذ محورين أولهما ترجمة معنى النص من لغة إلى لغة أخرى وهو ما متعارف عليه في مفهوم الترجمة أما المحور الآخر فيهتم بطريقة لفظ كلمات النص الأصلي (الفونيمات الصوتية) وتسمى ترجمة لفظية (سهام حسن، 2007). الترجمة اللفظية تساعد في قراءة نصوص اللغة بشكلها الصحيح ناهيك عن التمكن من اللغة من خلال سرعة التعلم والنطق السليم لأصواتها.

5. **النظام المقترح**

يتألف النظام المقترح من مرحلتين هما:

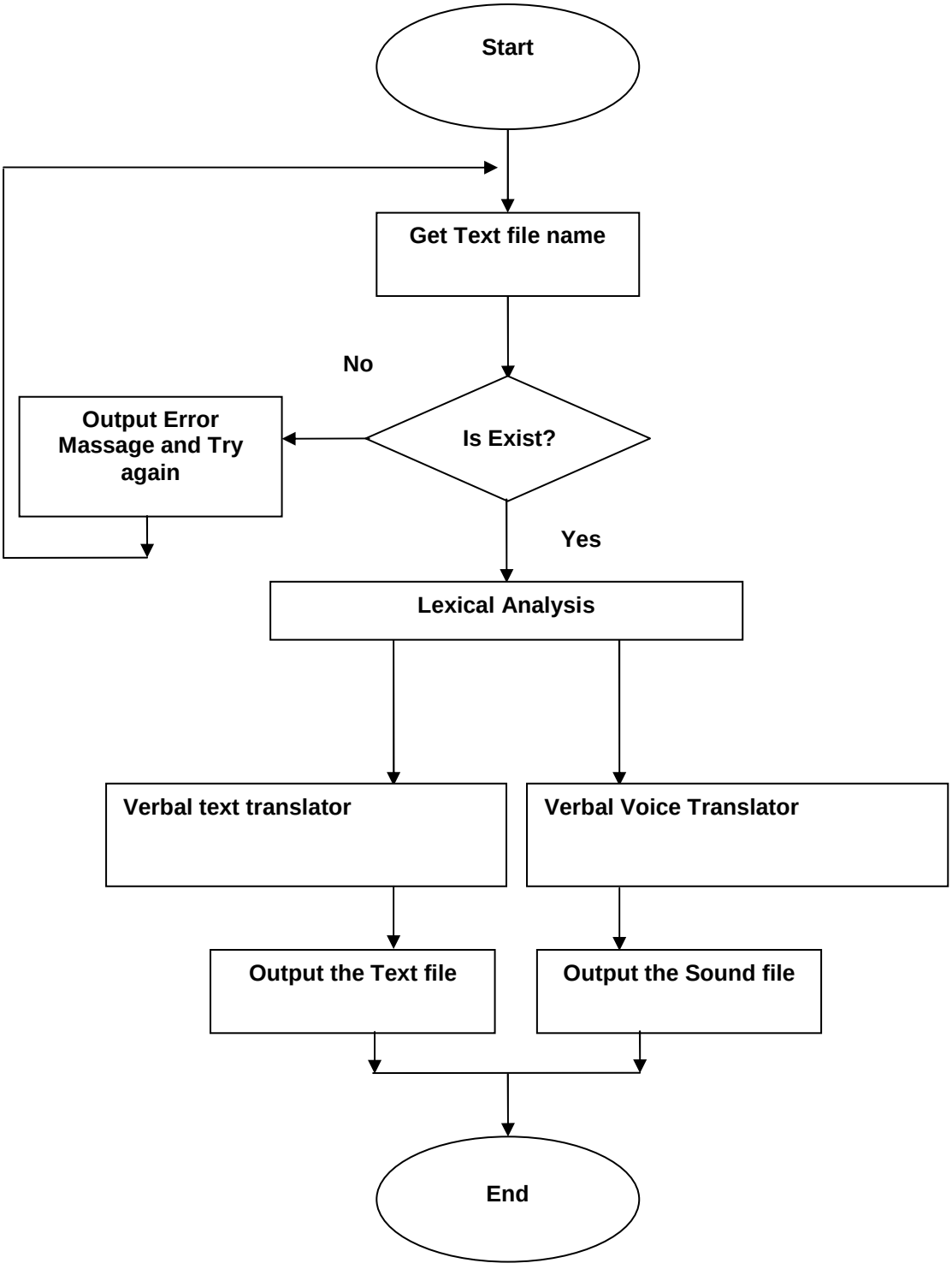
**المرحلة الأولى First Stage:** في هذه المرحلة قمنا بدراسة مستفيضة لأصوات اللغة الانكليزية من اجل حصر القواعد العامة للفظ ووضع قواعد أخرى حسب مارأيناه من خلال معالجة اللغة ومحاولة الاستفادة من النقاط المشتركة في لفظ الأصوات بغض النظر عن تصنيف الأصوات الانكليزية المذكور في الجدول(1) آنف الذكر، واستنادا إلى المعلومات التي استحصلت في هذه المرحلة تم وضع النموذج المعجمي الخاص بالترجمة اللفظية.

**المرحلة الثانية Second Stage:** في هذه المرحلة تم برمجة المسألة المقترحة باستخدام لغات C++ وMATLAB كانت معظم البرمجة بلغة C++ واستخدمنا MATLAB لمعالجة الملفات الصوتية لما تتمتع به من ميزات لمعالجة الصوت. والشكل (2) يوضح خوارزمية عمل البرنامج الرئيسية. إن مدخلات البرنامج هي عبارة عن اسم الملف النصي المطلوب ترجمته لفظيا مع مساره الخزني ويتم التأكد من وجود هذا الملف من عدمه من قبل البرنامج المقترح ثم يطبع رسالة خطأ في حالة عدم وجود الملف المطلوب أو كان هناك خطأ في إدخال اسم الملف أو مساره ويطلب من المستخدم إعادة إدخال الاسم الصحيح أما إذا وجد الملف تبدأ عملية المعالجة للملف بمرحلة تحليل المفردات ومن ثم يسلك البرنامج مسارين مختلفين كل مسار له مخرجات مختلفة. إن مخرجات المسار الأول (المترجم اللفظي الصوتي) هي عبارة عن ملف صوتي يتضمن المقاطع الصوتية لكل كلمات الملف المدخل، وملف نصي باللغة العربية يتضمن ترجمة لفظية حرفية لكل كلمات الملف المدخل كمخرجات للمسار الثاني(المترجم اللفظي النصي). لابد أن ننوه إن كلا المسارين يعملان في نفس الوقت ( آنيا).

### 1-5 محلل المفردات Lexical Analyzer

في هذه المرحلة يقوم محلل المفردات بتقطيع النص المدخل بعد فتح الملف وفحص حروفه فإذا كان هناك رمزا لا ينتمي للغة الانكليزية يقوم المحلل بحفظ هذه الرموز ومكان ورودها في النص في قائمة خاصة لتضاف لاحقا للنص بعد الانتهاء من الترجمة اللفظية دون تغيير كما في المعادلات الرياضية وبعض الرموز التي لا تتطلب ترجمة، وإلا يقوم المحلل بتكوين قائمة مرتبطة لكل الكلمات التي وردت في النص إذ يعتمد المحلل الفراغ بين كلمة وأخرى وستكون هذه القائمة مدخلات للمراحل اللاحقة. وكمثال على مخرجات هذه المرحلة لنأخذ هذه العبارة كمدخلات (Then supposes that X=Y*2) تكون المخرجات بشكل قائمة ["Then", "supposes", "that"]

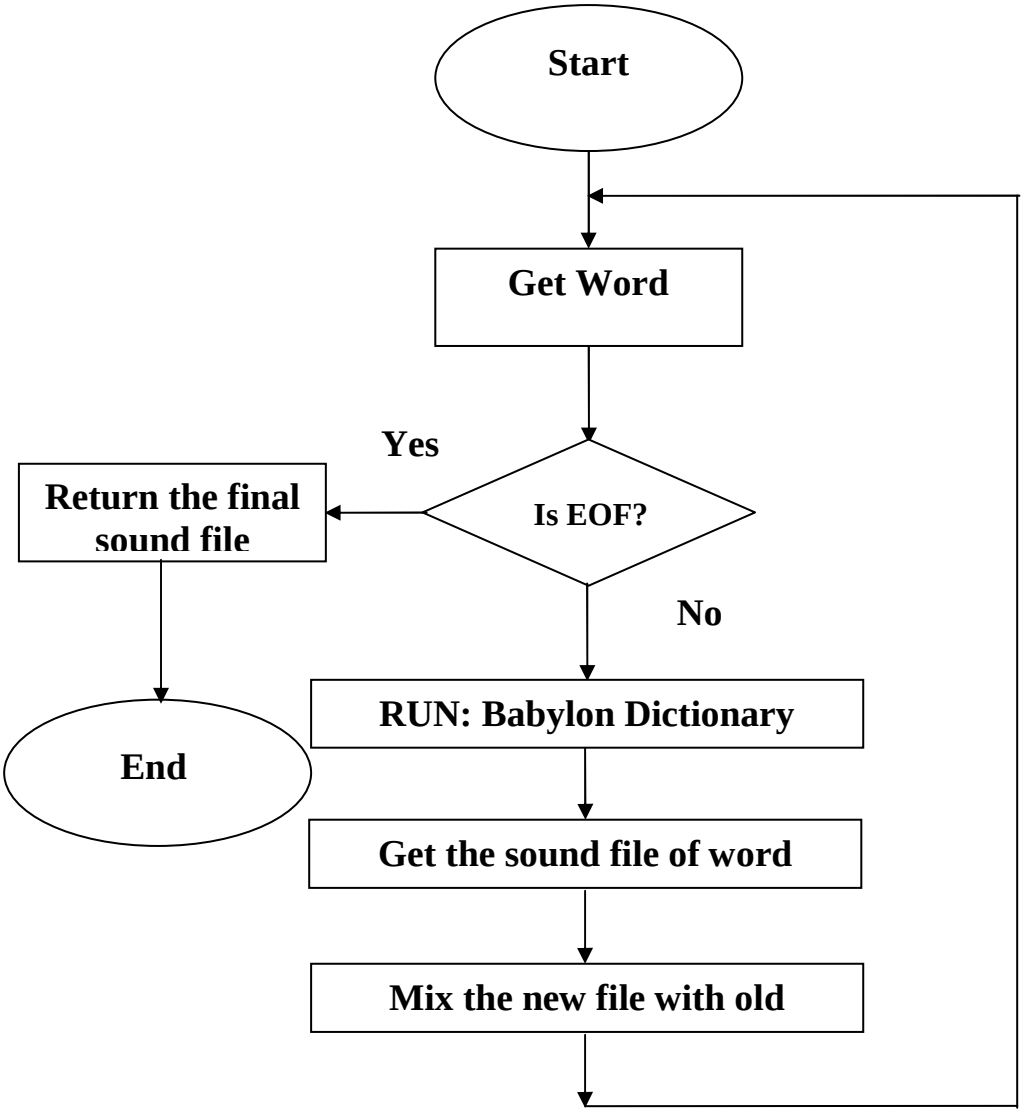
وقائمة للرموز الخاصة [('2',59) ('*',58) ('Y',57), ('=',56), ('X',55),] حيث تمثل الارقام (55، 56،...) تسلسل ورود الرمز نسبة للنص.




الشكل (2) يوضح خوارزمية النظام المقترح

## 2-5 المترجم اللفظي للصوت Verbal Voice Translator

يتم في هذه المرحلة معالجة القائمة الناتجة من محلل المفردات للحصول على ملف صوتي للملف المصدر وتم الاعتماد على برنامج القاموس الشهير(Babylon) والذي يعد من أهم القواميس الناطقة وأوسعها استخداما لما يتميز به من سهولة التعامل وتعدد المعاني والنطق السليم والواضح للمفردات حيث يعيد ملف صوتي لكل كلمة، والشكل (3) يوضح عمل خوارزمية المترجم اللفظي الصوتي. من خلال الشكل (3) نلاحظ إن خوارزمية المترجم اللفظي الصوتي تتقبل كمدخلات قائمة الكلمات الناتجة من محلل المفردات وتبدأ بمعالجة القائمة كلمة بعد أخرى في كل مرة يتم تشغيل القاموس لكلمة ومن ثم يعيد الملف الصوتي لهذه الكلمة، ليدمج مع الملف السابق وهكذا حتى نهاية القائمة يقوم المترجم اللفظي الصوتي بإعادة الملف الكلي إلى الخوارزمية الرئيسية. وهذا الملف يكون بصيغة (wav) ليتم تشغيله من قبل المستفيد فيما باستخدام أي برنامج تشغيل للأصوات مثل( Media player, real player, one real player,…etc)




الشكل (3) يوضح خوارزمية المترجم اللفظي للصوت

### 3-5 المترجم اللفظي للنص Verbal text translator

يقوم هذا البرنامج الفرعي بمعالجة الملف (المصدر) للحصول على ملف نصي باللغة العربية الذي يمثل ترجمة لفظية للملف المصدر بعدة خطوات كما في الشكل (4) الذي يمثل خوارزمية المترجم اللفظي النصي لكن قبل الخوض في تفاصيل هذه الخوارزمية لابد لنا من التكلم عن المعجم الذي تم اقتراحه كأساس تستند عليه هذه الخوارزمية في عملها.

### المعجم Lexicon

في مفهوم معالجة اللغات الطبيعية عند اقتراح أي نموذج معجمي لابد من الأخذ بعين الاعتبار أن يكون الخزن اقل ما يمكن وان يكون بشكل هيكل بياني سهل الوصول ومتعدد الاتجاهات، لذلك تم اختيار القوائم المرتبطة كهيكل بياني يتمتع بهذه المزايا المذكورة آنفا، كما إننا قللنا من الخزن عبر وضع مداخل مختصرة ذات قيم منطقية يمكن من خلالها معرفة صفة أو ميزة لطريقة لفظ كل حرف من حروف اللغة الانكليزية. الجدول (2) يوضح المعجم المقترح والذي يتكون من 26 مدخل (العمود الاول) بعدد حروف اللغة الانكليزية، نلاحظ من خلال الجدول إن هناك حروف لها صوت واحد وأخرى تكون متعددة الأصوات (العمود الثاني) لذلك يفتح المدخل عدة مداخل بعدد أصوات لفظ هذا الحرف كما في المدخل(a,d,..).

الجدول (2) يوضح المعجم المقترح

| Entry | matching | type | Mul | silent | Cond. | L. | V. | Anomalous words | Related entry |
|---|---|---|---|---|---|---|---|---|---|
| b | بِـ | 2 | N | Y | L,E,b | m | no | Debt, doubt, | no |
| p | بْـ | 1 | N | Y | no | no | no | Psychology,… | h |
| k | كْ | 1 | N | Y | L,A | n | no | … | no |
| l | … | … | … | … | … | … | … | … | … |
| n | … | … | … | … | … | … | … | … | … |
| m | … | … | … | … | … | … | … | … | … |
| r | … | … | … | … | … | … | … | … | … |
| w | … | … | … | … | … | … | … | … | … |
| . . . | . . . | . . . | . . . | . . . | . . . | . . . | . . | . . | . . |
| a | آ | 4 | Y | N | | | | …. | i,e,o,u,a |
| | أ | 3 | Y | N | | | | …. | i,e,o,u,a |
| | ايا | 3 | Y | N | L,b | e | | | i,e,o,u,a |
| | … | … | … | | | | | | i,e,o,u,a |
| . . . | . . . | . . . | . . . | . . . | . . . | . . . | . . | . . | . . |
| d | دَ | 2 | Y | N | E,V,A | no | 2,3,4 | …. | d |
| | اد | 2 | Y | N | E,V,A | no | D, t | …. | d |
| | تْ | 2 | Y | N | E,V,A | no | 1 | …. | d |
| | دَ | 2 | Y | Y | L,A | g | no | Handsome, … | d |
| . . . | . . . | . . . | . . . | . . . | . . . | . . . | . . | . . | . . |

أما العمود الثالث فيمثل نوع الصوت ويأخذ القيم(1,2,3,4) الذي تم اضافته الى الجدول (1) وذلك للمساعدة في عملية معالجة النص حيث من خلاله يتحدد الصوت اللاحق. العمود الرابع والخامس هما عبارة عن متغيرين منطقيين وقيمة اولهما تحدد هل ان للحرف (المدخل) اكثر من صوت ام لا، والاخر تكون قيمته نعم اذا كان هذا الحرف يأتي صامت في بعض الكلام ، اما الحقول الثلاث اللاحقة تعود لهذا الحقل فالسادس يمثل الشرط الذي يأتي لأجله الحرف صامت مثلا الحرف k لا يلفظ اذا جاء بعده الحرف n لذلك تكون قيم هذه الحقول الثلاث (L,A)(N)(NO) L تعني حرف وa تعني بعد الحرف الحالي حيث في مثالنا قيمة الحقل السابع هو n وهو الحرف الذي اذا اتى بعد الحرف الحالي(المدخل) لايلفظ هذا الحرف والحقل الثامن قيمته تمثل الصوت السابق او اللاحق الذي بشرط وجوده يكون الحرف صامت لأن في اللغة الانكليزية قواعد اللفظ تعتمد على الحرف او الصوت السابقين او اللاحقين. اما الحقل التاسع فيمثل قائمة الكلمات التي لايلفظ فيها هذا الحرف والتي تكون غير خاضعة لشروط معينة واسميناها قائمة الكلمات الشاذة، اما الحقل الاخير يحدد فيما اذا كان المدخل الحالي يرتبط بمداخل اخرى وماهي هذه المداخل.

**خوارزمية المترجم اللفظي للنص**

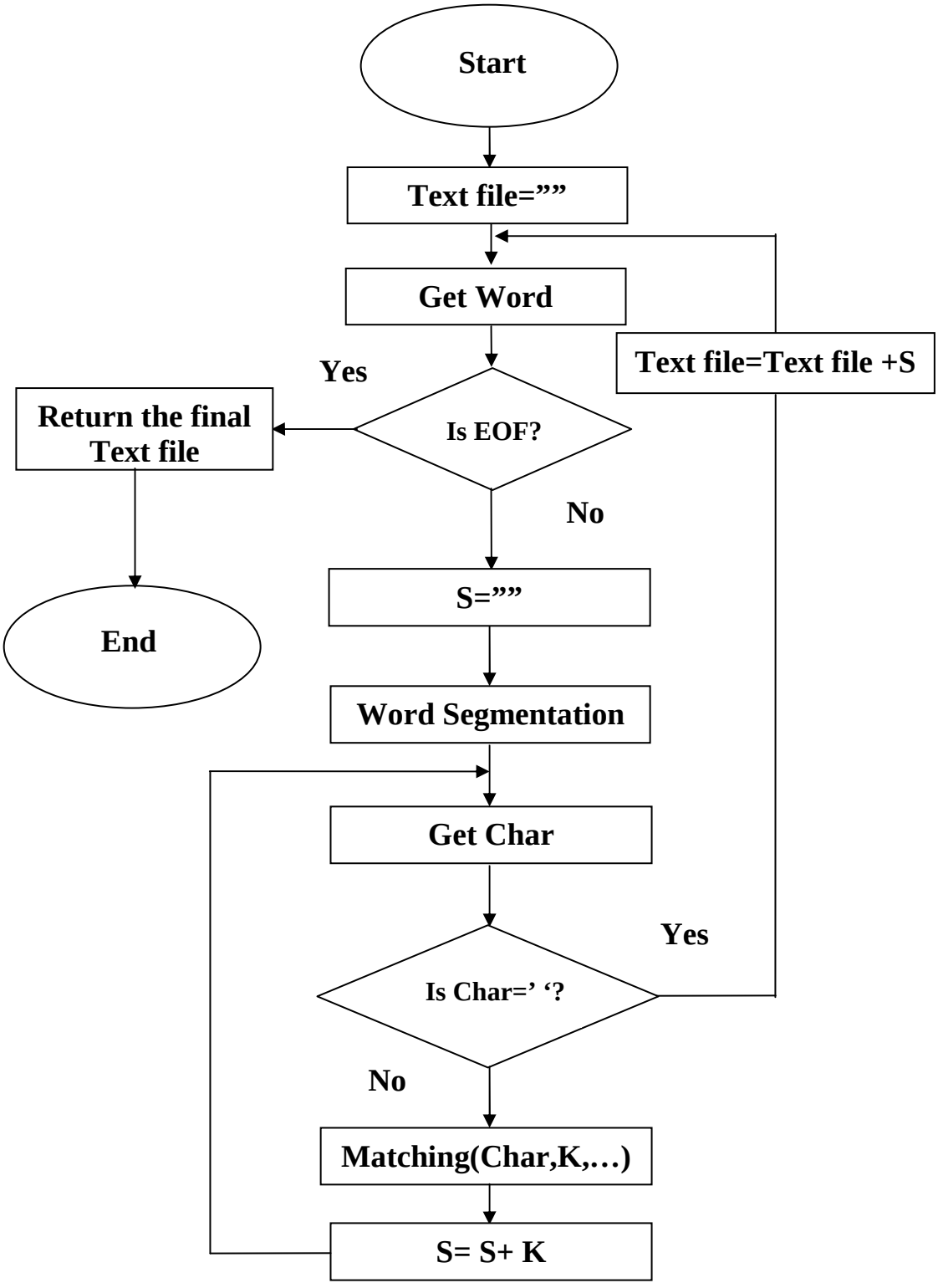
من خلال الشكل (4) نلاحظ ان خطوات الخوارزمية تبدأ بجعل الملف النصي الذي سيكون مخرجات لهذه المرحلة فارغ سيأخذ كلمة ومن ثم يتأكد هل وصلنا الى نهاية القائمة، اذا كان كذلك يقوم بإعادة الملف النصي الذي تم تجميعه في هذه المرحلة الى خوارزمية البرنامج الرئيسية والا تبدأ المعالجة حيث يقوم بتقطيع الكلمة المعطاة الى حروف( Word segmentation) ليتم معالجة حرف بعد اخر، يأخذ المترجم اللفظي النصي الحرف الحالي ويبحث في المعجم المقترح عن الحرف المطلوب(المدخل) حين يصل الى المدخل المطابق تبدأ عملية المطابقة (Matching) والتي يتم فيها عملية الحصول على الصوت المناسب لهذا الحرف اعتمادا على الشروط التي وضعناها في المعجم المقترح والتي ذكرناها آنفا، تعيد هذه المرحلة الصوت المقابل لهذا الحرف باللغة العربية اما في حالة وجد ان الحرف صامت فيعيد خيط رمزي فارغ، ويدمج الصوت القادم من مرحلة المطابقة مع بقية اصوات الكلمة وهكذا لحين نهاية الكلمة حيث يتم دمج المقطع الصوتي للكلمة التي تمت معالجتها الان مع بقية المقاطع الصوتية في الملف النصي وهكذا حتى نهاية الملف المصدر.

**6. الاستنتاجات**

إن النظام المقترح وسيلة جيدة لتعلم اللغة الانكليزية من خلال معرفة نطق الاصوات بشكل صحيح، يمكن ان يصير البرنامج المقترح كوسيلة منهجية لتعلم اللغة الانكليزية في المدارس والجامعات العراقية للتخلص من فقر اللغة الذي نعانيه.

ان النظام المقترح له القدرة على اعطاء نتائج ممتازة جدا بالنسبة للملف الصوتي ولا توجد نسبة خطأ تذكر، اما بالنسبة للملف النصي هناك نسبة خطأ بسيطة سببها حروف العلة غير المركبة لتعدد اصوات الحرف الواحد وعدم وجود قواعد محددة وان كانت متعددة.

في بداية العمل افترضنا ان الملف الصوتي المزعم اخراجه بصيغة (mp3) لكثرة البرامج الصوتية التي تتعامل مع هذا النوع من الملفات، لكن بعد ان قطعنا شوطا طويلا بالعمل واجهتنا صعوبة مفادها ان MATLAB لا تتعامل مع هذه الصيغة وانما يستطيع معالجة اربع انواع منها (wav, au) لذلك غيرنا نوع الملف الصوتي ليكون بصيغة wav وتعد هذه الصيغة من ابسط الصيغ وايضا تستعمل بكثرة ناهيك ان كل برامجيات تشغيل الاصوات قادرة على تشغيل هذه الصيغة. والمستخدم الذي يصر على استخدام الملف mp3 يمكنه تحويل الملف المعطى بصيغة wav الى صيغة mp3 من خلال احد برامجيات تحويل الصوت بكل يسر وسهولة.




الشكل (4) يوضح خوارزمية المترجم اللفظي للنص

## 7. العمل المستقبلي

ان النظام المقترح قد صمم للغة الانكليزية لكن بالامكان تطوير البرنامج ليساعد في تعلم اي لغة طبيعية اخرى من خلال حصر القواعد اللفظية لتلك اللغة وتغذيتها للبرنامج لتبقى طريقة المعالجة كما هي دون تغيير.

## 8. المصادر


1. Gibson, Edward &Pearlmutter, Neal J. "Constraints On Sentence Comprehension ", Cognitive Sciences, July, Volume 2 No. 7, 1998.
2. نبيل علي " اللغة العربية والحاسوب "، مؤسسة التعريب للنشر ، القاهرة ، 1988.
3. Charniak, Eugene & McDermott, Drew "Introduction to Artificial Intelligence" ,Addison-Wesley Publishing Company, 1988.
4. احمد قصير العاملي " متن الاجرومية ودروس في النحو "، دار مكتبة الهلال ، بيروت، 1985
5. Gazdar, Gerald & Mellish, Chris "Natural Language Processing in Prolog: An Introduction to Computational Linguistics ", California, USA, Addison-Wesley Publishing Company, 1989.
6. سهام حسن خيون " تصميم نظام ترجمة لفظية آلية باللغة الانكليزية للقرآن الكريم"، مجلة ابحاث البصرة (العلميات)، العدد 33، الجزء الثالث، 2007
7. سلمان داود الوسطي "التفاعل بين الانسان والآلة في الترجمة الحاسوبية" ، كلية الآداب ،الجامعة المستنصرية ،2002.
8. Michael Swan "Basic English Usage", Oxford New York, USA, Oxford University Press, 1995.

**Abstract:**

Many Arab researchers are facing a problem of reviewing of research, books, articles, …etc (written in English) because they are not able to read and translate them. There are many programs has the ability to translate from English into Arabic, but there is no program can assist for reading purpose, to enable the researcher from the language at all, to pronounce the letters in the right way, and finally to understand and assimilate what is stated in the texts.

Therefore, we decided to design a program has the ability to read texts in English, whether books, researches, or any text file, and then compose an audio file (with .wav format), and a text file (with .txt format) that contains a verbal translation of the English text in Arabic.
We Can be summarized the subsequence of the program by entering the text file name and its path by the user, then the program will analyze the text input to generate a list of all words, after this stage the program would take different tracks to emerge an audio file from the first track and the text file in Arabic from the second one.

**Track I:** Here, a sub-program named **verbal voice translator** uses an efficient search mechanism to search for each individual voice in the text (using one of the voiced dictionary), and stored the audio segment in the audio file, which the program make it during the collecting of the audio segment of the current word with the rest of the file and so on to the end of the list of words.

**Track II:** Here, we process the text according to the concept of **Natural Language Processing**, where another sub-program called **Verbal text translator** that translate English text into Arabic verbally and generate a text file for all words of the entered text. This processing is based on a lexicon which is designed to contain all the letters and sounds of English and the corresponding of Arabic.