مؤسسة النهضة العلمية

تعلم اللغة العربية بمساعدة الحاسب عن طريق الكلمات المتشابهة

أكرم محمد زكي و شيماء سنان محمود
الجامعة الإسلامية العالمية بماليزيا
akramzeki@iium.edu.my

الملخص: يتحدث الناس حول العالم بالعديد من اللغات المختلفة والتي تمتد بعض منها إلى مئات أو الآف السنين، ونشأت هذه اللغات بطرق مختلفة وانحدر بعض منها من لغات أخرى، ومع مرور الوقت توسع استخدام بعض هذه اللغات لتصبح لغات عالمية بينما تعرضت لغات أخرى إلى الاندثار. كما أن العديد من اللغات قد تعرضت للتحوير الذي يطرأ بشكل طبيعي على أي لغة مع مرور الوقت وبنفس الوقت تمت ولادة لغات جديدة من أعقاب لغات أخرى ولا شك ان هناك تشابهًا كبيرًا بين هذه اللغات الطبيعية، وتختلف نسبة التشابه ليشكل في بعض الأحيان ما يزيد عن نصف كلمات بعض اللغات والتي تعتبر مستلة من لغات أخرى. توسع استخدامات الحاسب الآلي في السنوات الماضية بشكل كبير حتى شملت استخداماته تعلم وتعليم اللغات الحية، ومن هذه اللغات حظيت اللغة الانجليزية بجانب كبير من هذا الاهتمام بينما لم تحظ اللغة العربية إلا على جانب قليل منه، وقد طورت في السنوات القليلة الماضية العديد من البرنامج والتطبيقات المستخدمة لتعليم اللغات ومنها اللغة العربية عن طريق الدروس والأمثلة والمسابقات والأسئلة والأجوبة والألعاب الحيوية التعليمية، بينما لا يوجد أي برنامج يقوم باكتشاف الكلمات المتشابهة في اللفظ والمعنى بين اللغة العربية واللغات الأخربشكل تلقائي ويستفيد من ذلك في تعليم اللغة، والتي تساعد هذه العملية على تسريع التعلم بشكل سهل وفعال وممتع، حيث ان المستخدم سيسهل عليه تذكر الكلمات ذات اللفظ المتشابه مع لغته الأم. يهدف هذا البحث إلى استخدام الحاسب وتقنية المعلومات لتسريع تعلم وتعليم اللغة العربية للمتحدثين بغيرها (اللغة الانجليزية أنموذجا في هذا البحث) وذلك عن طريق التركيز على الكلمات المتشابهة بين اللغة العربية واللغة الانجليزية، سيتمكن المتعلم عن طريق البرنامج من إنشاء وإضافة الدروس وسيتم استخراج الكلمات المتشابهة آلياً عن طريق الحاسب، سيتم استخدام المنهج التحليلي لدراسة وتحليل الطرق والخوارزميات المستخدمة لإيجاد التشابه بين اللغات المختلفة ومن ثم تطوير نظام لإيجاد التشابه بين اللغة العربية والانجليزية. وتعد هذه الأنظمة إحدى طرائق معالجة اللغات الطبيعية.

الكلمات الجوهرية: الكلمات المتشابهة، خوارزمية الكلمات المتشابهة، تعلم اللغة بالحاسب.

1. المقدمة:

يعتبر تعليم اللغات الطبيعية من الأمور الأساسية في عصرنا الحديث حيث تبرز حاجة الناس المستمرة إلى اكتشاف العلوم المنشورة باللغات الأخرى بالإضافة إلى الحاجة الكبيرة للتواصل بين الشعوب المختلفة خاصة وأن العالم قد أصبح قرية صغيرة يمكن للمرء التواصل بشكل مباشر مع جميع انحاء العالم بالصوت والصورة في أي وقت ومن أي مكان. لقد ساهمت تقنية المعلومات بشكل كبير في تعلم وتعليم اللغات الحية وحظيت اللغة الانجليزية بالجانب الأوفر من هذا الاهتمام بينما تختلف حظوظ اللغات الأخرى بشكل متفاوت بما فيها اللغة العربية. يهدف هذا البحث إلى استخدام الحاسب وتقنية المعلومات لتسريع تعلم وتعليم اللغة العربية للمتحدثين بغيرها وذلك عن طريق التركيز على الكلمات المتشابهة بين اللغة العربية واللغات الأخرى، وسيتم استخراج الكلمات المتشابهة آلياً عن طريق الحاسب.

يقدم هذا القسم خلفية موجزة عن الموضوع الذي أجراه هذا البحث، ويركز على أهمية اللغة العربية من جوانب مختلفة حيث أنها إحدى أكثر اللغات شيوعًا التي يتحدث بها الناس ويرغبون في تعلمها .اللغة العربية هي اللغة الأم الخامسة الأكثر شيوعًا في العالم. بالإضافة إلى ذلك فهي اللغة الرسمية لخمس وعشرين دولة عربية، وكذلك اللغة الرسمية لمنظمة مؤتمر التعاون الإسلامي وجامعة الدول العربية وغيرها من المنظمات العالمية.

كان للغة العربية تأثيراً كبيراً على اللغات الأخرى، وخاصة في المفردات وقد كان الأثر البالغ في تلك البلدان التي يقطنها المسلمون. فاللغة العربية هي المصدر الرئيسي للمفردات للعديد من اللغات مثل الأمهرية والبنغالية والهندية والإندونيسية والكازاخستانية والكردية والقرغيزية والمالاوية والباشتو والفارسية والبنجابية والسندية والصومالية والسواحيلية والتغرينية والتركية والتركمانية والأوردو والأويغور والأوزبكية بالإضافة إلى لغات أخرى. على سبيل المثال، تُستخدم الكلمة العربية للكتاب (kitāb كتاب) في معظم اللغات المذكورة أما خارج نطاق العالم الإسلامي فقد كان للغة الإسبانية (على سبيل المثال) أكبر المفردات العربية وذلك بسبب الحكم الإسلامي في شبه الجزيرة الأيبيرية من عام 711م حتى 1492م المعروفة بالأندلس. ومع هذا فإن كلمة كتاب لا تقابلها كلمة مشابهة باللغة الإسبانية.

ورغم أهمية اللغة العربية عالمياً، إلا أن كثير من الناطقين بغيرها يواجهون صعوبة في تعلمها كلغة ثانية. ومن ثم، فمن المستحسن أن يستخدم المعلمون تقنية المعلومات لتعزيز تعليمهم للغة العربية، حيث إن استخدام التقنية والوسائط المتعددة يحفز المتعلمين ويعزز أدائهم في الفصل. ورغم ازدياد استخدام الحاسب في تعلم اللغات بشكل عام إلا أن استخدام الحاسب لتعلم اللغة العربية ليس بهذه البساطة. حيث أن أحد التحديات التي تواجه تعلم اللغة العربية بمساعدة الحاسب هو أنها لا تستخدم الحروف اللاتينية

التي حظيت بكثير من الاهتمام في العقود الأخيرة. ومن أهم الفروق بين اللغة العربية واللغات اللاتينية أن عدد الحروف العربية بشكل عام أكثر من غيرها وطريقة النطق مختلفة، كما أنها تكتب من اليمين إلى الشمال، وبحروف ملتصقة ويتغير شكل الحرف حسب موقعه في الكلمة، ومن خصائص اللغة العربية أنها تكتب كما تلفظ، وكثير من الكلمات تعود إلى نفس الجذر وهذا ما يميز اللغة العربية عن غيرها، كما أنها تحوي الحركات وهذه ميزة أخرى للغة العربية، بالإضافة إلى أن الجمل الفعلية في العادة تبدأ بالفعل.

ورغم هذه الميزات العديدة للغة العربية إلا أن تقنيات الحاسب في تعليم اللغة العربية لا يزال في مرحلة التطوير[1]. إن التحدي يكمن في الافتقار إلى موارد الحاسب لتعليم اللغة العربية[2].

لتعزيز تعلم اللغة العربية، يمكن استخدام طريقة التعلم عبر الكلمات المتشابهةcognates (المتشابهة في المعنى والنطق). إن استخدام الكلمات المستعارة من لغة معينة للمساعدة في تعلم اللغة الأخرى (كما هو الحال في اللغة العربية) تعد استراتيجية فعالةلتوسيع المفردات. علاوة على ذلك، فإن لفت انتباه المتعلمين إلى التشابه اللغوي يمكن أن يساعد في توسيع مفرداتهم وتحسين فهمهم[3]. وبالتالي، تكمن أهمية هذا العمل البحثي في إيجاد بيئة إبداعية لتعلم اللغة العربية. حيث يهدف هذا البحث إلى مساعدة المتعلم على التغلب على تحديات وصعوبات تعلم اللغة العربية من خلال استخدام مزيج من منهجية تعليم اللغة العربية عبر الحاسبمع استخدام التشابه اللغوي بين اللغات المختلفة.

ومن المؤمل أن يعمل البحث على تعريف الطلاب الناطقين باللغة الإنجليزية بالتشابه اللغوي في اللغتين العربية والإنجليزية، مما يمكنهم من فهم الارتباط بين هاتين اللغتين وبالتالي مساعدتهم في تعلم اللغة العربية. علاوة على ذلك، يهدف البحث إلى تعزيز أنظمة التعلم الإلكترونية العربية الحالية من خلال إضافة إستراتيجية التعلم الذاتي وزيادة الوعي حولها.

---

[1]Ghalib, M., Hassan, H., Sahrir, M. S., Daud, N. M., Mohin, M., & Ayu, M. A. (2014). Developing Icollect Mobile (Android and Iphone-Based) Application for Arabic Language Teaching. *Middle-East Journal of Scientific Research*, *20*, 21–27

[2]Suliman, I., Mukhtar, A., &Sahrir, M. S. (2013). Issues and Challenges of Using CALL in Learning Arabic : Considerations for Content Development. *American International Journal of Social Science*, 2(5), 76–88.

[3]Goldberg, Y., Elhadad, M. (2008). Identification of Transliterated Foreign Words in Hebrew Script. In: Gelbukh, A. (eds) Computational Linguistics and Intelligent Text Processing. CICLing 2008. *Lecture Notes in Computer Science*, vol 4919. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-78135-6_40

2. مشكلة البحث:

تعد الدراسات في مجال تعليم وتعلم اللغة العربية باستخدام الحاسب قليلة نسبياً، كذلك الحال بالنسبة للدراسات التي ركزت على التشابه اللغوي بين اللغات المختلفة، أما الدراسات التي جمعت بين الكلمات المتشابهة والتعليم عن طريق الحاسب فهي غائبة عن الحضور، وهو ما يحاول هذا البحث التركيز عليه، وتحديداً للناطقين باللغة الانجليزية.فنحن بحاجة إلى تطوير تطبيقات لتعليم اللغة العربية عن طريق التشابه اللغوي بين اللغات المختلفة (وخاصة بين اللغة العربية والانجليزية)، بالإضافة إلى دراسة تأثير ذلك على المعلمين والمتعلمين.

3. أسئلة البحث وأهدافه:

يمكن تلخيص أهم أسئلة البحث بما يلي:

- ما المقصود من تعلم اللغة العربية عن طريق الكلمات المتشابهة مع اللغات الأخرى في اللفظ والمعنى؟
- كيف يمكن أن يساعد الحاسب من تسريع عملية تعلم اللغة العربية للناطقين بغيرها؟
- هل يمكن للمتعلم أن يبدأ بتعلم اللغة العربية عن طريق الكلمات المتشابهة مع لغته الأم؟
- كيف يمكن للحاسب من إيجاد الكلمات المتشابهة بين اللغة العربية وغيرها بشكل تلقائي؟

ومن أبرز أهداف هذا البحث ما يلي:

- تشجيع متعلمي اللغة العربية على تعلمها بشكل سهل وسريع وفعال عن طريق التركيز على الكلمات المتشابهة مع لغتهم الأم.
- تطوير برنامج (تطبيق إلكتروني) لاكتشاف الكلمات المتشابهة في المعنى واللفظ بين اللغة العربية واللغة الانجليزية.
- بناء قاعدة بيانات للكلمات المتشابهة في المعنى واللفظ في كل من اللغة العربية والانجليزية.
- مساعدة المتعلم على إنشاء الدروس التعليمية وسيقوم البرنامج بدوره من تحديد الكلمات المتشابهة في هذه الدروس بين اللغة العربية والانجليزية (باللفظ والمعنى) مع إيجاد نسبة التشابه.

4. تعلُم اللغة بمساعدة الحاسب:

في السنوات السابقة تم استخدام الحاسبوتقنياته كأداة في جميع مراحل التعلم، فظهرت آلاف البرمجيات والتطبيقات التي سهلت عملية التعلم، بما في ذلك تعلم اللغات. في مجال اللغة، فإن تعلم اللغة بمساعدة

الحاسب هو نهج لاستخدام الحاسب كأداة لتعلم اللغة وتعليمها[1]. حيث أن تعلم اللغة بمساعدة الحاسب يجعل التعلم أسهل وأكثر مرونة وأكثر كفاءة،كما هو الحال في العديد من مواقع الإنترنت وحزم البرمجيات والتطبيقات المتوفرة على أجهزة الهواتف الذكية والتي يمكن الوصول إليها بسهولة. وتعد الموارد والمعدات التعليمية الوفيرة التي يوفرها تعليم اللغة بمساعدة الحاسب منهجية جديدة لاستبدال طرق التعلم التقليدية، لأنها تعمل في بيئات أكثر مرونة ولديها خيارات أفضل فيما يتعلق باحتياجات المتعلم والمعلم على حد سواء، واهتماماتهم وقدراتهم. كما أن أدوات الوسائط المتعددة وعناصرها المتعددة تشكل إضافة نوعية لأنظمة تعليم اللغة عبر الحاسب، لأنها تتيح التكامل مع الرسومات والفيديو والصوت بالإضافة إلى النص.

كثير من طرائق تعلم اللغات عن طريق الحاسب تتم باستخدام ما يعرف بتقنيات معالجة اللغات الطبيعية. حيث تتضمن هذه التقنيات استخدام البيانات اللغوية لتدريب نماذج الذكاء الصناعي على فهم وإنتاج النصوص. يتم ذلك من خلال تغذية النماذج بملايين النصوص بمختلف اللغات والمواضيع، مما يمكنها من اكتساب فهم عميق للبنية اللغوية والمعاني. حيث يمكن للحاسب توليد نصوص، وترجمة النصوص، وفحص النحو والإملاء، والإجابة على الأسئلة، والتفاعل مع المستخدم بشكل طبيعي. ويتطلب تدريب النماذج جهداً كبيراً لمعالجة وفهم البيانات اللغوية المتنوعة، وتحسين أدائها عبر عمليات التحسين المتكررة.ويمكن تلخيص أهم فوائد تعلم اللغة عن طريق الحاسب بما يلي:

1. ترجمة فعالة وسريعة من لغة إلى أخرى: يمكن للحاسب تقديم خدمات ترجمة آلية دقيقة وسريعة بين مختلف اللغات، مما يسهم في تسهيل التواصل بين الناس والتفاعل فيما بينهم.
2. تحليل النصوص واستخراج المعلومات: تستطيع تقنيات الحاسب تحليل النصوص الكبيرة واستخراج المعلومات الهامة منها، مثل مراجعات المستخدمين، ومقالات الأخبار، والمحادثات عبر وسائل التواصل الاجتماعي.
3. المساعدة الشخصية للمستخدمين: يمكن للحاسبتقديممساعدة شخصية ذكية للمستخدمين كالإجابة على الأسئلة، وتقديم النصائح، وتنفيذ مهام معينة بناءً على الأوامر الصوتية أو النصية.
4. تحسين تعلم اللغة: يمكن استخدام التطبيقات والمنصات التعليمية لتعليم اللغة باستخدام تقنيات تعلم الآلة، حيث يمكنها تقديم تدريب مخصص ومتجدد باستمرار وفقاً لاحتياجات كل فرد.

[1]Abu Naba’ih, A., Hussain, J., Al-omari, A., &Shdeifat, S. (2009). The Effect of Computer Assisted Language Learning in Teaching English Grammar on the Achievement of Secondary Students in Jordan. *The International Arab Journal of Information Technology*, *6*, 431–439.

5. توفير المحتوى وإنتاج النصوص: يمكن للحاسب إنتاج محتوى نصي متنوع، مثل المقالات، والتقارير، والمحتوى الإعلاني، مما يخدم العديد من الصناعات والمجالات.
6. دعم اتخاذ القرارات: يمكن عن طريق تحليل البيانات اللغوية الكبيرة من تقديم مساعدة في اتخاذ قرارات بشكل أفضل واستنادًا إلى الأدلة المتاحة.
7. تواصل إنساني مع الحاسب: يمكن للنماذج اللغوية الذكية تحسين تجربة التواصل بين الإنسان والحاسب، حيث يمكن للأفراد التفاعل بطريقة طبيعية مشابهة للتواصل بين البشر أنفسهم.

تعتمد فوائد تعلم اللغة عن طريق الحاسب على استخدام التقنيات الحديثة والتقدم في مجال معالجة اللغة الطبيعية، ويمكنها أن تحقق تأثير إيجابي على العديد من جوانب الحياة اليومية والأعمال. بالإضافة إلى ما سبق فهناك تقنيات أخرى بجانب معالجة اللغة الطبيعية التي يمكن استخدامها لتعلم اللغة. بعض هذه التقنيات تركز على جوانب مختلفة من التعلم والتفاعل مع اللغة. ومن ذلك ما يلي:

1. تعلم اللغة من خلال التفاعل الاجتماعي: هذه التقنيات تهدف إلى تحسين التفاعل الاجتماعي وتعلم اللغة من خلال المحادثات الحقيقية مع أشخاص متحدثين باللغة المستهدفة.
2. تعلم اللغة عبر المحادثات الصوتية والدروس الصوتية: تطبيقات ومنصات تقدم دروس صوتية ومحادثات مع متحدثين بلغتهم الأم، مما يساهم في تطوير مهارات الاستماع والنطق.
3. تعلم اللغة من خلال الألعاب والتطبيقات التفاعلية: تقنيات تعليم اللغة المبتكرة تستخدم الألعاب والتطبيقات التفاعلية لجذب الأفراد وتشجيعهم على تعلم اللغة بطريقة ممتعة ومشوقة.
4. تعلم اللغة من خلال الدروس المكتوبة والمقروءة: استخدام الدروس المكتوبة والنصوص المقروءة يمكن أن يساهم في تحسين مهارات القراءة وفهم النصوص باللغة المستهدفة.
5. تعلم اللغة عن طريق الترجمة والتفاعل مع النصوص: استخدام تطبيقات الترجمة والتفاعل مع النصوص المختلفة يمكن أن يساهم في تطوير مفردات وفهم البنية اللغوية.
6. تعلم لغة الإشارة عبر الحاسب: تقنيات تعلم لغة الإشارة عبر الحاسب بل وتعليم الحاسب كيفية التواصل مع ذوي الاحتياجات الخاصة حيث تستخدم الإشارات البصرية بدلاً من الكلمات المنطوقة أو المكتوبة.

تختلف هذه التقنيات في الأهدافوالمنهجية، وتتيح تنوع الاختيارات عبر وسائل مختلفة في تعليم اللغةلتحسين مهارات المتعلمين بشكل أفضل.

5. التعلم بالتشابه اللغوي:

الكلمات المتشابهة هي الكلمات التي لها نفس المعنى والنطق في لغتين مختلفتين[1]. وهذا التشابه لا يدل بالضرورة على أن هناك استعارة من لغة إلى أخرى ولا يدل على أن الأصل لكلا اللغتين هو واحد، ولهذا السبب فقد أبدى أحد الأبحاث[2] اعتراضه على من يحددون أوجه التشابه بين اللغات على أنها بسبب الميراث اللغويمن لغة إلى أخرى.

وهناك من يذكر أسبابًا أخرى لوجود التشابه بين اللغات، على سبيل المثال عن طريق الحدث العارض (مصادفة)، أو المحاكاة الصوتية، أو السمات العامة والسمات الشائعة من الناحية النمطية. تنص احدى الأبحاث[3] على أن تعلمالكلمات المتشابهةcognates يمكن أن يساعد في تسهيل تعلم لغة أجنبية أو لغة ثانية لأن المتعلمين يحتفظون بالمعلومات اللغوية للغة الأولى. هناك اتفاق واسع على أن استخدام الكلمات المتشابهة يساعد في تسهيل عملية تعلم اللغات، خاصةً إذا كانت الكلمات المتشابهة ذات تردد مرتفع نسبيًا.

وعند الحديث عن التشابه اللفظي بين اللغات (التشابه في النطق)، يجب التطرق إلى موضوع الكتابة الصوتية (التحويل الصوتي من العربية إلى الإنجليزية) حيث توجد مناهج وطرق مختلفة للترجمة الصوتية للغة العربية. وهي تختلف في طريقة معالجتها لمشاكل الترجمة التقليدية للنص اللاتيني إلى اللغة العربية المكتوبة والمنطوقة. ومن أمثلة هذه المشكلات رموز الصوتيات العربية التي لا توجد في اللغة الإنجليزية أو اللغات الأوروبية الأخرى؛ ومن أمثلة ذلك حروف العلة في اللغة العربية والتي يتم نطقها بشكل موحد في اللغة العربية لكن كتابتها بالحروف اللاتينية متعددة فعلى سبيل المثال نجد أن كلمة "مسلم" قد تكتب بعدة أشكال في الحروف اللاتينية مثل Muslim/Moslem، وكذلك الحال في كلمة "محمد" حيث نجد عدة صيغ لكتابتها في الحروف اللاتينية مثل:Mohammed/Muhammad/Mohamed، وغير ذلك من الكلمات.

[1]Sherif, T., Kondrak, G., (2007) "Bootstrapping a stochastic transducer for Arabic-English transliteration extraction". In: Proc. of *ACL*.

[2]Tadayonifar, M., Entezari, M., Valizadeh, M.: The effects of computer-assisted L1 and L2 textual and audio glosses on vocabulary learning and reading comprehension across different learning styles. J. Lang. Educ. 7(2), 223–242 (2021).

[3]Hara, Y. (2011). *The Use Of Loanwords In English Vocabulary Learning*. Ball State University, Muncie, Indiana.

ويمكن اتباع بعض الإرشادات العامة للترجمة الصوتية التي وضعتها المجلة الدولية لدراسات الشرق الأوسط[1]، وغيرها من الإرشادات، ويمكن اختصارها بما يلي:

- في حال وجود الكلمة العربية مترجمة ومكتوبة حسب النطق في القواميس والمعاجم اللغوية فيتم اعتمادها.
- في حال عدم وجود الكلمة العربية مكتوبة حسب النطق فيتم كتابتها بالحروف اللاتينية، مع اعتماد ما يلي:
  - عدم ترجمة أسماء الأشخاص مثل "جمال" و "كريم" والعناوين والمصطلحات الفنية وكذلك أسماء المنظماتوالجمعيات، وإنما تكتب حسب النطق.
  - ترجمة أسماء المدن والبلدان حسب الترجمة المعتمدة، مثل "الأردن" يكتب Jordan، وفي حال عدم وجود ترجمة للمدينة مثل مدينة "بعلبك" فتكتب حسب النطق.
  - عدم إهمال علامات التشكيل / الحركات باللغة العربية ويجب كتابتها بالحروف اللاتينية حسب التهجئة.
  - استخدام العلامة العليا "’" Apostrophe في الحروف مثل حرف العين والهمزة، فعلى سبيل المثال تكتب كلمة "شريعة" كما يلي: Shari'ah.
  - مراعاة الجمع للكلمات، فعلى سبيل المثال إذا تم ترجمة كلمة "مفتي" بكلمة Mufti فيكون جمعها بإضافة حرف s في آخرها Muftis.

6. الأعمال ذات الصلة

يناقش هذا القسم بعض الدراسات السابقة المتعلقة بهذا البحث. هناك أبحاث محدودة جدًا تم إجراؤها حول استخدام نظام التعلم عن طريق الحاسب باستخدام الكلمات المتشابهة للغة العربية والإنجليزية. على الجانب الآخر، هناك العديد من الدراسات التي تستخدم الكلمات المتشابهة للغات أخرى. في احدى الدراسات[2]، وصفوا طريقة للتعرف التلقائي على التشابه اللغوي في النصوص أحادية اللغة باستخدام الترجمة الآلية. فقد اقترحوا أخذ نص مُدخلًا باللغة الأصلية وبإستخدام ترجمة آلية إحصائية يتم إنشاء نص موازٍ محاذٍ للكلمات في اللغة المطلوبة. تم بعد ذلك تطبيق مقياس قوي لمسافة السلسلة، مسافة Jaro–Winkler في هذه الحالة، على أزواج الكلمات المتوافقة لاكتشاف الكلمات المتشابهة المحتملة.

[1]IJMES, *International Journal of Middle East Studies*, "IJMES Translation & Transliteration Guide". Retrieved 2016 December

[2]Malmasi, Shervin and Dras, Mark, (2015) "Cognate Identification using Machine Translation", *Proceedings of the Australasian Language Technology Workshop (ALTA*), 138-141. Sydney, Australia.

استنادًا إلى الافتراض بأن التغييرات اللغوية للكلمات تتبع قواعد معينة، اقتُرحت[1] طريقة للكشف التلقائي عن التشابه اللغوي باستخدام طريقة المحاذاة الإملائية والتي أثبتت أنها ذات صلة بمحاذاة التسلسل في علم الأحياء الحسابي. تم استخدام المتشابهات المتوافقة كميزات لخوارزميات التعلم الآلي من أجل استنتاج قواعد التغييرات اللغوية التي تمر بها الكلمات عند إدخال لغات جديدة والتمييز بين الكلمات المتشابهة وغير المتشابهة.

استكشفت دراسة أخرى[2] دور علم الأصوات في معالجة الكلمات العربية عند المتحدثين الأصليين للغة الإنجليزية الذين يتعلمون اللغة العربية. استخدمت كلمات إنجليزية مدمجة صوتيًا في الكلمات العربية في ثلاث نقاط مختلفة في الكلمة: البداية، الوسط، والنهاية، لتحديد دور علم الأصوات المنفصل عن دور الدلالات. تدعم النتائج بشكل عام أهمية علم الأصوات في المعالجة العقلية للغات متعددة، والتي يمكن دمجها مع النتائج الأخرى في أبحاث اللغة لتكملة برامج تعلم اللغة.

تم تقديم طريقة خاضعة للإشراف الحر[3] لتحديد الهوية بدون سياق للأسماء الأجنبية المترجمة والكلمات المستعارة في النص العبري. الطريقة إحصائية بحتة ولا تتطلب استخدام أية معاجم أو أدوات تحليل لغوي للغات الأصلية (العبرية، في هذه الدراسة). كما أنه لا يتطلب أي بيانات مشروحة يدويًا للتدريب. تم الإبلاغ عن نتائج الدقة / الاستدعاء 82/80 لمجموعة مكونة من 4044 كلمة فريدة تحتوي على 368 كلمة أجنبية.

استخدمت احدى الدراسات[4] أمثلة أولية ونصًا إنجليزيًا / عربيًا محاذاة للجملة لتعلم دالة المسافة للسلسلة ثنائية اللغة بشكل مشترك واستخراج أزواج مترجمة. بينما يهدف هذا العمل إلى المواءمة الكاملة، كانت مهمتهم هي تحديد المرشحات المترجمة فقط. يمكن أن يساعد تحديد المرشحات للترجمة الصوتية على المحاذاة الكاملة عن طريق التخفيف من الحاجة إلى نص محاذي.

---

[1]Ciobanu, Alina Maria and Liviu P. Dinu (2014) "Automatic Detection of Cognates Using Orthographic Alignment"*Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics* (Short Papers), pages 99–105,Baltimore, Maryland, USA, June 23-25 2014.

[2]Biller, Alia K. (2010) "Translation Recognition in Learners of Arabic", *Inquiry journal*, Volume 11, 2010.

[3]Goldberg, Y., Elhadad, M. (2008). Identification of Transliterated Foreign Words in Hebrew Script. In: Gelbukh, A. (eds) Computational Linguistics and Intelligent Text Processing. CICLing 2008. *Lecture Notes in Computer Science*, vol 4919. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-78135-6_40

[4]Sherif, T., Kondrak, G., (2007) "Bootstrapping a stochastic transducer for Arabic-English transliteration extraction". In: Proc. of *ACL*

في هذه الدراسة[1]، تم التركيز على استخدام الترجمة الصوتية للعربية لتحسين نتائج طريقة محاذاة الكلمات القائمة على اللغويات من مجموعة نصوص متوازية. من ناحية تستخدم هذه الطريقة معجمًا ثنائي اللغة، وكيانات مسماة، وما شابه ذلك، وعلامات نحوية لمحاذاة الكلمات المفردة. ومن ناحية أخرى، تستخدم علاقات التبعية النحوية لمحاذاة الكلمات المركبة. لقد قاموا بتقييم كلمة محاذاة لدمج الترجمة الصوتية للغة العربية باستخدام طريقتين:تقييم يدوي لجودة المحاذاة وتقييم لتأثير هذه المحاذاة على جودة الترجمة باستخدام نظام موسى للترجمة الآلية الإحصائية. تظهر النتائج التي تم الحصول عليها أن الترجمة الصوتية للغة العربية تعمل على تحسين جودة كل من المحاذاة والترجمة.

الترجمة الآلية المستندة إلى المثال[2] (EBMT) شكل من أشكال الترجمة الآلية التي تستخدم مجموعة كبيرة من جمل الأمثلة المترجمة مسبقًا لإنشاء ترجمة لجملة جديدة. عادة لا يحتوي النظام في مجموعته على الجملة الكاملة المراد ترجمتها. بدلاً من ذلك، يطابق النظام الكلمات والعبارات الصغيرة ويجمعها معًا بمساعدة نموذج اللغة المستهدفة. على الرغم من أن EBMT يستخدم طرقًا إحصائية، إلا أنه يختلف عن الترجمة الآلية الإحصائية (SMT). تستشير EBMT مجموعة الترجمات الخاصة بها في وقت التشغيل، بينما تعالج SMT المجموعة مسبقًا لحساب احتمال حدوث كلمة أو عبارة كترجمات، ولكن بعد ذلك تستخدم هذه الاحتمالات فقط في وقت التشغيل.

وفي خلاصة هذه الدراسات السابقة يمكن القول، على الرغم من وجود كثير من الدراسات في هذا الموضوع، إلا أنه لم يتم تطوير نظام أو برنامج للتعليم الإلكتروني لتعلم اللغة باستخدام معرفة التشابه اللغوي في أي دراسة. ومن الواضح أن هناك نقصًا في هذا الميدان، كما أن هناك حاجة لمعرفة تأثير تطبيق مثل هذه الأنظمة على المعلمين والمتعلمين وقياس كفاءة أداءها.

[1]Houda Saadane, OuafaBenterki, Nasredine Semmar, and Christian Fluhr. 2012. Using Arabic Transliteration to Improve Word Alignment from French- Arabic Parallel Corpora. In *Fourth Workshop on Computational Approaches to Arabic-Script-based Languages*, pages 38–46, San Diego, California, USA

[2]Aaron B., Philips, Violetta-Sforza (2012) "Arabic-to-English Example Based Machine Translation Using Context-Insensitive Morphological Analysis", *4th International conference on Arabic language Processing*.

7. منهج البحث:

يوضح هذا القسم أهم مراحل العمل التي يتبعها هذا البحث، كما هو موضح أدناه:

- المرحلة الأولى وهي مرحلة المسح الأدبي للبحوث المهتمة بتعليم اللغات عن طريق الكلمات المتشابهة.
  - دراسة مفهوم تعلم التشابه اللغوي وتوظيف منهج تعلم اللغة بمساعدة الحاسب في تسهيل تعلم اللغة خاصة باللغة العربية من خلال مراجعة الدراسات السابقة ذات الصلة.
- المرحلة الثانية وهي جمع البيانات من القواميس والمعاجم المعتمدة باللغتين (العربية والإنجليزية) وتضمينها في قاعدة البيانات.
  - إدخال عينة مختارة من الكلمات العربية مع المعنى المقابل لها باللغة الانجليزية.
  - إعادة كتابة الكلمات أعلاه وفق اللفظ المناسب لكل كلمة Transliteration.
- المرحلة الثالثة وهي مرحلة بناء وتطوير نظام إلكتروني (برنامج) لإيجاد الكلمات المتشابهة
  - استخدام المنهج التحليلي لدراسة وتحليل الطرق السابقة والخوارزميات المستخدمة لإيجاد التشابه بين الكلمات المختلفة.
  - تطوير نموذج أولي (خوارزمية) لقياس مقدار التشابه بين أي كلمتين مختلفتين بين النسب 0 – 100%.
- المرحلة الرابعة وهي مرحلة التنفيذ
  - تنفيذ البرنامج واستخراج قواعد بيانات للكلمات المتشابهة بين اللغتين مع إيجاد نسبة التشابه بين هذه الكلمات.
  - إنشاء وإضافة الدروس عن طريق المستخدم ليتم تحديد الكلمات المتشابهة فيها.
- المرحلة الخامسة وهي مرحلة التقييم
  - اختبار النظام وتصحيح الأخطاء
  - عرض النتائج على خبراء لتقييم أداء البرنامج.

الدوال المستخدمة لتحديد التشابه:

يسمى تحديد التشابه بين سلسلتين "مسافة ليفينشتين". تمت تسمية مسافة ليفينشتاين على اسم عالم الرياضيات فلاديمير ليفينشتاين، الذي طور الخوارزمية في عام 1965[1]. وتم تطويرها برمجيًا بعد ذلك وظهرت العديد من الدوال (أو الوظائف (functions)) التي تُستخدم لتحديد التشابه بين سلسلتين / كلمتين. ومن أهم هذه الدوال هي()similar_text و ()levenshtein.

[1]Gosselin, D., Kokoska, D., & Easterbrook, R. (2011). *PHP Programming With MYSQL* (2nd ed.). Boston, MA: Course Technology, Cengage Learning.

- **الدالة الأولى (() similar_text)**

ترجع الدالة similar_text () إلى عدد الأحرف التي تشترك فيها سلسلتان.
يقبل وسيطتا سلسلة تمثلان القيم التي يجب مقارنتها. على سبيل المثال، سيؤدي تطبيق الدالة (أو الوظيفة) على وسيطتي السلسلة "car" و "can" إلى إرجاع النتيجة "2" نظرًا لوجود حرفين مشتركين بين السلسلتين. يسمح باستخدام وسيطة ثالثة بديلة تُستخدم لتخزين القواسم المشتركة بين السلسلتين كنسبة مئوية[1]. ويمكن استخدام هذه الدالة كما يلي:

similar_text(string1,string2,percent)

وتحوي هذه الدالة على ثلاثة عناصر، كما يلي:

- **العنصر الأول** وهي السلسلة الأولى هي الترجمة الصوتية للكلمة العربية المخزنة في قاعدة البيانات (وفي حال لم تكن متوفرة في قاعدة البيانات فيتم استخراجها بشكل تلقائي).
- **والعنصر الثاني** وهي السلسلة الثانية في الدالة وهي المعنى الإنجليزي للسلسلة / الكلمة العربية الأولى المخزنة في قاعدة البيانات.
- **والعنصر الثالث** وهي النسبة المئوية، وهي وسيطة غير معروفة سابقاً وسيتم حسابها في البرنامج، حيث سيتم تمريرها في الدالة لحساب التشابه بالنسبة المئوية بين السلسلة الأولى (التحويل الصوتي للغة العربية) والسلسلة الثانية (الكلمة الإنجليزية). وعلى المستخدم تحديد النسبة المئوية المطلوبة والتي تعتبر مقبولة بالنسبة له (بعض المستخدمين يفضل أن تكون نسبة التشابه 75% أو أكثر وآخرين يقبلون بنسبة تشابه 50% أو أكثر).

وبعد تحديد جميع العناصر يمكن عندها كتابة البرنامج وتطوير الخوارزمية وهي سلسلة من التعليمات الواضحة لحل مشكلة، أي للحصول على المخرجات المطلوبة لأي مدخلات مشروعة في فترة زمنية محدودة، ويمكن تحويل الخوارزمية إلى مخطط انسيابي (مجموعة من الأشكال الهندسية المتصلة التي تحتوي على وصف لخطوات عمل الخوارزمية) ويوضح الشكل 1 (الملحق 1) مخطط انسيابي لعمل الدالة أو الوظيفة المقترحة.

[1]Tatroe, K., MacIntyre, P., & Lerdorf, R. (2013). Programming PHP (3rd ed.). North, Sebastopol, CA: *O'Reilly Media*, Inc.

- **الدالة الثانية (levenshtein ()):**

تُرجع الدالة ()levenshtein عدد الأحرف المختلفة بين السلسلة الأولى (الكلمة الأولى) والسلسلة الثانية (الكلمة الثانية) بناء على المقارنة بينهما حرفاً تلو الآخر وإيجاد التماثل والاختلاف [1]. فعلى سبيل المثال، سيؤدي تطبيق الدالة على الكلمتين "car" و "can" إلى إرجاع النتيجة إلى "1" نظرًا لوجود حرف واحد فقط مختلفًا بينهما. وتكتب الدالة بهذا الشكل:

levenshtein(string1,string2,insert,replace,delete)

وكما هو واضح حيث يوجد في الدالة ثلاثة عناصر:

العنصر الأول وهي السلسلة الأولى هي الترجمة الصوتية للكلمة العربية المخزنة في قاعدة البيانات (وفي حال لم تكن متوفرة في قاعدة البيانات فيتم استخراجها بشكل تلقائي).

والعنصر الثاني وهي السلسلة الثانية في الدالة وهي المعنى الإنجليزي للسلسلة / الكلمة العربية الأولى المخزنة في قاعدة البيانات.

والعنصر الثالث وهي تختلف عن الدالة السابقة حيث لن يتم حساب نسبة التشابه، بين العنصر الأول والثاني وإنما حساب عدد الأحرف التي يجب استبدالها أو حذفها إو إضافتها لتصبح كلا الكلمتين متماثلتين،

ويجب حساب طول السلسلة الأولى والثانية للعثور على السلسلة الأطول، ومن ثم تنفيذ المعادلة الآتية لحساب نسبة التشابه.

((طول السلسلة –عدد الأحرف المختلفة) / الكلمة الأطول) * 100

حيث تكون النتيجة النهائية هي قيمة النسبة المئوية.

وهنا أيضاً على المستخدم تحديد النسبة المئوية المطلوبة والتي تعتبر مقبولة بالنسبة له، حيث سيتم تمرير جميع البيانات المخزنة في قاعدة بيانات في الدالة، فإذا كان التشابه في النسبة المئوية بين "السلسلة الأولى" (التحويل الصوتي للغة العربية) و "السلسلة الثانية" (الإنجليزية) يساوي أو أكبرمن قيمة "النسبة المئوية" المعطاة ، فستحدد الدالة الكلمة على أنها كلمة متشابهة، وخلاف ذلك لا يتم تحديد الكلمة ضمن المتشابهات. ولتوضيح ذلك، يمكن الإطلاع على مثال توضيحي من خلال الشكل 2 (الملحق 2).

[1]Gosselin, D., Kokoska, D., & Easterbrook, R. (2011). *PHP Programming With MYSQL* (2nd ed.). Boston, MA: Course Technology, Cengage Learning.

8. النتائج الأولية:

بعد أن يحدد المستخدم الحد الأقصى لمعدل الخطأ المقبول (أو عدة معدلات اعتمادًا على خطورة الأخطاء) والحد الأدنى للنسبة المئوية المقبولة، يمكن الرجوع إلى الجدول 1 (الملحق 3) لمعرفة النسبة المئوية المقبولة بشكل عام[1]. ونجد أدناه بعض الأمثلة لبعض الكلمات:

- مثال 1:
  - الكلمة العربية: ليمون
  - الكلمة العربية بالحروف الاتينية: līmūn
  - الترجمة الانجليزية: lemon
  - نسبة التشابه: 3/5 = 60%
- مثال 2:
  - الكلمة العربية: سكر
  - الكلمة العربية بالحروف الاتينية: Sukkar
  - الترجمة الانجليزية: Sugar
  - نسبة التشابه: 4/6 = 66.6%
- مثال 3:
  - الكلمة العربية: نظير
  - الكلمة العربية بالحروف الاتينية: naẓīr
  - الترجمة الانجليزية: nadir
  - نسبة التشابه: 3/5 = 60%
- مثال 4:
  - الكلمة العربية: سلطان
  - الكلمة العربية بالحروف الاتينية: sultān
  - الترجمة الانجليزية: sultan
  - نسبة التشابه: 5/6 = 83%
- مثال 5:
  - الكلمة العربية: عفريت
  - الكلمة العربية بالحروف اللاتينية: ʿifrīt
  - الترجمة الانجليزية: afrit
  - نسبة التشابه: ʿifrīt – afrit : 3/5 = 60%

[1]John Lehrberger, Laurent Bourbeau "Machine Translation: Linguistic characteristics of MT systems and general methodology of evaluation", *John Benjamins Publishing*, retrieved 2016 December

تم تطبيق الخورازمية المقترحة على مجموعة من الكلمات باللغة العربية واللغة الانجليزية واستخراج النتائج كما مبين في الجدول 2 (الملحق 4).

9. مناقشة:

توسعاستخداماتالحاسب الآلي في السنوات الماضيةبشكلكبيرحتشملتاستخداماتهتعلم وتعليماللغات الحية، ومن هذه اللغات حظيت اللغة الانجليزية بجانب كبير من هذا الاهتمام بينما لم تحظ اللغة العربية إلا بجانب قليل منه. يحاول هذا البحث إيجاد طريقة فعالة لتعليم اللغات الأجنبية للناطقين بغيرها. ركز هذا البحث على اللغة العربية كلغة تعلم للناطقين بلغة أخرى (اللغة الانجليزية على سبيل المثال كما في هذا البحث) عن طريق تطوير برنامج يقوم باكتشاف الكلمات المتشابهة في اللفظ والتي لها نفس المعنى بطريقة تلقائية آلية. يقوم البرنامج باستخراج الكلمات المتشابهة لفظاً مع وضع نسبة التشابه لكل كلمة، والتي يمكن استخدام النتائج لتسريع عملية التعلم حيث ان المستخدم سيسهل عليه تذكر الكلمات ذات اللفظ المتشابه مع لغته الأم. يمكن للمستخدم إنشاء الدروس وإضافتها للبرنامج، وسيقوم البرنامج بدوره من تحديد الكلمات المتشابهة في هذه الدروس ليتمكن المتعلم من معرفتها والتركيز عليها. كما يمكن للبرنامج بعد تطويره من اضافة لغات أخرى بطريقة فعالة ويسيرة. قد يستفيد الكثير من الأشخاص والعديد من القطاعات التعليمية من النظام المقترح ومنهم على سبيل المثال ما يلي:

- متعلمو اللغات الأفراد
- معاهد تعليم اللغات
- مدرسو اللغات
- المدارس والجامعات
- الشركات التجارية المهتمة بتعليم اللغات
- العاملين في قطاع الترجمة

10. الخاتمة

رغم وجود العديد من التطبيقات والبرنامج المستخدمة لتعليم اللغة العربية للناطقين بغيرها عن طريق الدروس والأمثلة والمسابقات والأسئلة والأجوبة والألعاب التعليمية، بينما لا يوجد أي برنامج يقوم باكتشاف الكلمات المتشابهة في اللفظ والمعنى مع اللغة الأم، يهدف هذا البحث إلى تطوير برنامج يعمل على تحديد الكلمات المتشابهة بين اللغة العربية مع لغة المتعلم الأم (اللغة الانجليزية على سبيل المثال) بطريقة إلكترونية تلقائية ويحدد البرنامج نسبة التشابه عن طريق خوارزمية مبتكرة. كما يمكن للمتعلم من إضافة الدروس في البرنامج وسيقوم البرنامج بدوره من تحديد

ومن الآفاق المستقبلية لهذا البحث أن يتم كتابة الكلمات بالحروف العربية وليس الآتينية وإضافة لغات أخرى، كما أن عملية التشابه يجب أن تتطور لتشمل التشابه الصوتي بين الحروف بناء على مخارج الحروف فعلى سبيل المثال حرف الحاء والهاء فيهما تشابه، لكن الدوال المستخدمة تعتبرها مختلفين تماماً مثلهما مثل أي حرفين مختلفين كحرف الحاء والقاف. فاعتماد مخارج الحروف في حساب نسبة التشابه يؤدي إلى دقة أكبر في النتائج.

11. المراجع والمصادر

1. Aaron B., Philips, Violetta-Sforza (2012) "Arabic-to-English Example Based Machine Translation Using Context-Insensitive Morphological Analysis", 4th International conference on Arabic language Processing.
2. Abas, A., Abusahyon, A. S., &Sahad, M. S. (2019). Semantically Divergent Meaning of Arabic Loanwords in the Malay Language. *Abas, NA Abusahyon, AS &Sahad, MN*, 29-30.
3. Abu Naba'ih, A., Hussain, J., Al-omari, A., &Shdeifat, S. (2009). The Effect of Computer Assisted Language Learning in Teaching English Grammar on the Achievement of Secondary Students in Jordan. *The International Arab Journal of Information Technology*, *6*, 431–439.
4. Abufanas, O. (2013). Computer Aided Language Learning System for Arabic for Second Language Learners. *International Journal of Social, Management, Economics and Business Engineering*, *7*(12), 1942–1946.
5. Al-bazeli, A., Alreshidi, A. N. G., Naimie, Z., & Halili, S. H. (2014). Malaysian Students Learning The Arabic Language In Arab Schools In Malaysia: A Future Study. *The Online Journal of New Horizons in Education*, *4*(4), 119–126.
6. Al-Zubaidi, K. O. (2018). The Main Characteristics of Arabic Borrowed Words in Bahasa Melayu. *AWEJ for Translation & Literary Studies*, *2*(4).
7. Baugh, A. C., & Cable, T. (2012). *A History of the English Language* (6th ed.). Longman.
8. Bekele, T. A. (2010). Motivation and satisfaction in internet-supported learning environments: A review. *Educational Technology and Society*.
9. Biller, Alia K. (2010) "Translation Recognition in Learners of Arabic", Inquiry journal, Volume 11, 2010.
10. Chih-Ming, C., Ying-You, L.: Developing a computer-mediated communication competence forecasting model based on learning behavior features 1, 100004 (2020). https://doi.org/10.1016/j.caeai.2020.100004
11. Ciobanu, Alina Maria and Liviu P. Dinu (2014) "Automatic Detection of Cognates Using Orthographic Alignment"*Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics* (Short Papers), pages 99–105, Baltimore, Maryland, USA, June 23-25 2014.
12. Dina, A., & Ciornei, S.-I. (2013). The Advantages and Disadvantages of Computer Assisted Language Learning and Teaching for Foreign Languages. *Procedia - Social and Behavioral Sciences*, *76*, 248–252. doi:10.1016/j.sbspro.2013.04.107
13. Ghalib, M., Hassan, H., Sahrir, M. S., Daud, N. M., Mohin, M., & Ayu, M. A. (2014). Developing Icollect Mobile (Android and Iphone-Based) Application for Arabic Language Teaching. *Middle-East Journal of Scientific Research*, *20*, 21–27. doi:10.5829/idosi.mejsr.2014.20.lcl.203

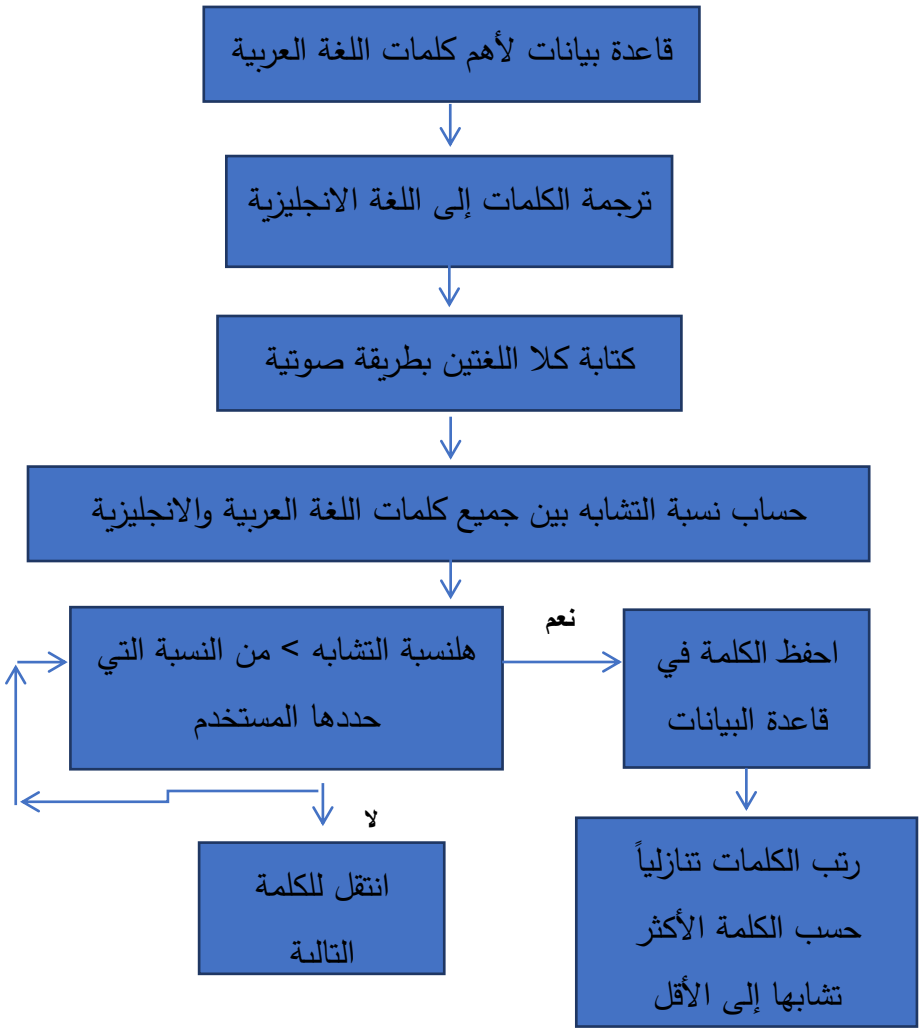
14. Goldberg, Y., Elhadad, M. (2008). Identification of Transliterated Foreign Words in Hebrew Script. In: Gelbukh, A. (eds) Computational Linguistics and Intelligent Text Processing. CICLing 2008. Lecture Notes in Computer Science, vol 4919. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-78135-6_40
15. Gosselin, D., Kokoska, D., & Easterbrook, R. (2011). *PHP Programming With MYSQL* (2nd ed.). Boston, MA: Course Technology, Cengage Learning.
16. Guarda, M. (2012). Computer-mediated-communication and foreign language education. *Journal of E-Learning and Knowledge Society*, *8*, 15–27.
17. Hamida, N. (2012). Learning Styles and Strategies in Second Language Learning. *Ragam JurnalPengembanganHumaniora*, *12*(1), 8–14.
18. Hara, Y. (2011). *The Use Of Loanwords In English Vocabulary Learning*. BALL STATE UNIVERSITY, MUNCIE, INDIANA.
19. https://ijmes.chass.ncsu.edu/IJMES_Translation_and_Transliteration_Guide.htm
20. Hung, B.P., Pham, A.T.D., Purohit, P. (2022). Computer Mediated Communication in Second Language Education. In: Sharma, R., Sharma, D. (eds) New Trends and Applications in Internet of Things (IoT) and Big Data Analytics. Intelligent Systems Reference Library, vol 221. Springer, Cham. https://doi.org/10.1007/978-3-030-99329-0_4
21. Hussain, A. (2012). *Textual Similarity*. Bachelor thesis, Technical University of Denmark.
22. Ibrahim, M. N., Bilmas, S. N., Babiker, A., &Idroas, M. (2013). A framework of an online self-based learning for teaching arabic as second language (TASL). In *Proceedings of International Conference on Computational Intelligence, Modelling and Simulation* (pp. 255–260). doi:10.1109/CIMSim.2013.48
23. Ibrahim, R. (2006). Do Languages with Cognate Relationships have Advantages in Second Language Acquisition?.*The Linguistics Journal*, *1*(3), 66-96.
24. IJMES, International Journal of Middle East Studies, "IJMES Translation & Transliteration Guide". Retrieved 2016 December
25. Jethro, O. O., Grace, A. M., & Thomas, A. K. (2012). E-Learning and its effects on teaching and learning in a global age. *International Journal of Academic Research in Business and Social Sciences*, *2*, 73–78.
26. John Lehrberger, Laurent Bourbeau "Machine Translation: Linguistic characteristics of MT systems and general methodology of evaluation", John Benjamins Publishing, retrieved 2016 December
27. Levitin, A. (2012). *Introduction to the design & analysis of algorithms* (3rd ed.). Boston: Pearson.
28. Mahmoudi, E., Samad, A. B. A., & Razak, N. Z. B. A. (2012). Attitude and Students' Performance in Computer Assisted English Language Learning (CAELL) for Learning Vocabulary. *Procedia - Social and Behavioral Sciences*, *66*, 489–498. doi:10.1016/j.sbspro.2012.11.293
29. Malmasi, Shervin and Dras, Mark, (2015) "Cognate Identification using Machine Translation", Proceedings of the Australasian Language Technology Workshop (ALTA), 138-141. Sydney, Australia.
30. Nachoua, H. (2012). Computer-Assisted Language Learning for Improving Students' Listening Skill. *Procedia - Social and Behavioral Sciences*, *69*(Iceepsy), 1150–1159. doi:10.1016/j.sbspro.2012.12.045
31. Otwinowska-Kasztelanic, A. (2012). Plurilingualism and Polish Teenage Learners of English. *Lingvarvm Arena*, *3*, 37–52.

32. Pham, H.T.P.: Computer-mediated and face-to-face peer feedback: student feedback and revision in EFL writing. Comput. Assist. Lang. Learn. 34(4) (2021). https://doi.org/10.1080/09588221.2020.1868530
33. Pudaruth, S., &Jannnoo, N. B. (2010). An E-Learning tool for the self-study of mathematics for the CPE examination. *World Academy of Science, Engineering and Technology*, *46*, 427–432.
34. S Norfaizah, & Muhammad. (2011). *Perbualanharian :bahasa Arab, Melayu, Inggeris*. Selangor, Malaysia: AR-Risalah.
35. Houda Saadane, OuafaBenterki, Nasredine Semmar, and Christian Fluhr. 2012. Using Arabic Transliteration to Improve Word Alignment from French- Arabic Parallel Corpora. In *Fourth Workshop on Computational Approaches to Arabic-Script-based Languages*, pages 38–46, San Diego, California, USA.
36. Sahrir, M. S., & Yusri, G. (2012). Online Vocabulary Games For Teaching And Learning Arabic. *GEMA OnlineTM Journal of Language Studies*, *12*(September), 961–977.
37. Sahrir, M. S., Yahaya, M. F., & Nasir, M. S. (2013). EZ-Arabic for Children: A Virtual Learning Resource Tool for Malaysian Primary Schools. *Procedia - Social and Behavioral Sciences*, *90*(InCULT 2012), 396–404. doi:10.1016/j.sbspro.2013.07.108
38. Salkind, N. J. (2010). *Encyclopedia of Research Design*. SAGE Publications.
39. Schepens, J., Dijkstra, T., Grootjen, F., & van Heuven, W. J. B. (2013). Cross-Language Distributions of High Frequency and Phonetically Similar Cognates. *PLoS ONE*, *8*. doi:10.1371/journal.pone.0063006
40. Shehab, Roaya et. al. "Arabic-Malay Cognates As A Computer Assisted Language Learning", 3rd International Conference on Advanced Computer Science Applications and Technologies (ACSAT 2014) 29-30 Dec 2014 Amman, Jordan
41. Sherif, T., Kondrak, G., (2007) "Bootstrapping a stochastic transducer for Arabic-English transliteration extraction". In: Proc. of ACL.
42. Soleimani, H., & Molla Esmaeili, M. (2012). The Effect of Explicit Instruction of Clustering New Words on Vocabulary Learning of Iranian Intermediate EFL Learners through Hyperlinks. International Journal of Applied Linguistics & English Literature, 1(2), 112–126. doi:10.7575/ijalel.v.1n.2p.112
43. Stoimenova, S. T. (2010). Evaluation of Web Based Information Systems : Users ' Informing Criteria. *Issues in Informing Science and Information Technology*, *7*.
44. Suliman, I., Mukhtar, A., &Sahrir, M. S. (2013). Issues and Challenges of Using CALL in Learning Arabic : Considerations for Content Development. *American International Journal of Social Science*, *2*(5), 76–88.
45. Sunderman, G., & Schwartz, A. I. (2008). Using Cognates to Investigate Cross-Language Competition in Second Language Processing. TESOL Quarterly, 42(3), 527-536.
46. Tadayonifar, M., Entezari, M., Valizadeh, M.: The effects of computer-assisted L1 and L2 textual and audio glosses on vocabulary learning and reading comprehension across different learning styles. J. Lang. Educ. 7(2), 223–242 (2021). https://doi.org/10.17323/jle.2021.11020
47. Tatroe, K., MacIntyre, P., & Lerdorf, R. (2013). Programming PHP (3rd ed.). North, Sebastopol, CA: O'Reilly Media, Inc.
48. Tavakoli, M., & Gerami, E. (2013). The Effect of Keyword and Pictorial Methods on EFL Learners ' Vocabulary Learning and Retention. *Porta Linguarum*, *19*, 299–316.
49. Trentman, E. (2011). L2 Arabic Dialect Comprehension: Empirical Evidence for the transfer of Familiar Dialect Knowledge to Unfamiliar Dialects. L2 Journal, 3(1).
50. Vu, N.N., Hung, B.P., Van, N.T.T., Lien, N.T.H.: Theoretical and instrucional aspects of using multimedia resources in language education: a cognitive view. In: Kumar R.,

Sharma R., Pattnaik P.K. (eds.) Multimedia Technologies in the Internet of Things Environment, vol. 2, pp. 165–194. Springer (2021)

51. Wang, Z., Han, F.: Developing English language learners' oral production with a digital game-based mobile application. PLoS ONE 16(1), e0232671 (2021). https://doi.org/10.1371/journal.pone.0232671
52. Webber, A. B. (2010). *Modern Programming Languages: A Practical Introduction* (2nd ed.). Franklin, Beedle & Associates Inc.
53. Whiteman, M. (2012). Learning Spoken Arabic through Computer-Assisted Instruction: An integrated approach.
54. Wong, L. L. C., & Nunan, D. (2011). The learning styles and strategies of effective language learners. *System*. doi:10.1016/j.system.2011.05.004

الملحق 1



الشكل 1: المخطط الإنسيابي لخطوات عمل الخوارزمية المقترحة

الملحق 2

1) (الجبر, algebra)
a l g e b r (transliteration of الجبر)
a l g e b r a (the English word)

Levenshtein = 1 (an 'a' is added at the end)
Similarity percentage = ((the length of longest string – levenshtein) / the length of longest string) * 100
= ((7 – 1) / 7) * 100 ≈ 85.7 %

2) (دون, Down)
Doon (transliteration of دون)
Down (the English word)

Levenshtein = 1 (one letter is difference / one change is needed)
Similarity percentage = ((the length of longest string – levenshtein) / the length of longest string) * 100
= ((4 – 1) / 4) * 100 = 75%

3) (مدرسة, school)
m a d r a s a h (transliteration of مدرسة)
s c h o o l (the English word)
Levenshtein = 8 (all the letters need to be replaced or deleted)

Similarity percentage = ((the length of longest string – levenshtein) / the length of longest string) * 100
= ((8 – 8) / 8) * 100 = 0%

الشكل 2: أمثلة لإيجاد نسبة التشابه باستخدام دالة levenshtein()

الملحق 3

الجدول 1: فئات نسبة التشابه المقبولة

الفئات	نسب التشابه
متطابق	100%
متشابه	90% أكثر من
مقبول	75% - 90%
وسط	50% - 75%
ضعيف	40% - 50%
غير متشابه	40% أقل من

الملحق 4

الجدول 2: مجموعات أولية من النتائج

الكلمات باللغة (العربية)	الترجمة الصوتية	الترجمة (باللغة الانجليزية)	نسبة التشابه
قابل	Qabbil (Adj.)	Able	50 %
جرة	Jarra	Jar	60 %
المتراس	mitrass	Mattress	62.5 %
جدول	Jedule	schedule	62.5 %
(يشفط) شفط	shafit	shaft	83 %
صلد	sald	solid	60 %
الذيل	thail	tail	80 %
إذن	ethen	then	80 %
العنبر	Anbar	Amber	60 %
بدوي	bedoui	bedouin	85.7 %
البرميل	barmel	barrel	83 %
قانون	qanon	canon	80 %
قول	qaol	Call	50 %
قناة	qana	canal	60 %
كوب	cuop	Cup	75 %

12. السيرة الذاتية لمؤلف

يعمل الدكتور أكرم محمد زكي أستاذاً في قسم أنظمة المعلومات في كلية تقنية المعلومات والاتصالات في الجامعة الإسلامية العالمية بماليزيا، وله العديد من المؤلفات والبحوث والكتب المنشورة في مجال تقنية المعلومات وما تقدمه من خدمات خاصة في خدمة اللغة العربية والتطبيقات الإسلامية، أشرف على ما يزيد عن 20 رسالة ماجستير ودكتوراه، أدار عدداً من المشاريع البحثية المدعومة من الجامعة أو من وزارة التعليم في ماليزيا أو من خارج ماليزيا. حصل على العديد من الجوائز وتقلد العديد من المناصب داخل الجامعة الإسلامية منها منسق الدراسات العليا وثم رئيس البحث في كلية تقنية العلومات والاتصالات ثم نائب مدير المركز الإسلامي في الجامعة الإسلامية العالمية بماليزيا ثم مديراً بالوكالة للمركز. شارك في العديد من اللجان العلمية داخل وخارج الجامعة الإسلامية العالمية بماليزيا، وعمل كممتحن للعديد من البرامج الأكاديمية، ومحكم في عدد من المسابقات الدولية. للدكتور أكرم أكثر من 10 كتب منشورة، وما يزيد عن 150 بحثاً منشوراً في محرك البحث العلمي (سكوبس)، وما يزيد عن 45 فصلاً في كتب علمية، وشارك في العديد من المؤتمرات المحلية والدولية.

13. الخلاصة باللغة الانجليزية

Computer Assisted Language Learning for Arabic using Cognates

Akram M. Zeki and Shayma'a Senan Mahmood

International Islamic University Malaysia
akramzeki@iium.edu.my

Abstract: People around the world speak many different languages, some of which extend back to hundreds or thousands of years, and these languages arose in different ways and some of them descended from other languages. Many languages have undergone modification that occurs naturally in any language with time. At the same time, new languages were born from other languages, and for sure there are large similarities between these languages, and the similarity rate varies. Therefore, it sometimes can be more than half of the language's words, which are considered derived from other languages. The use of computers in the past years has expanded greatly to be used in various aspects of human life including learning and teaching living languages. Among these languages, English language received a large part of this attention, while the Arabic language received only a small part of it. In the past few years, many programs and applications used to teach Languages, including the Arabic language, through lessons, examples, competitions, questions and answers in addition to educational games. While there is no program that detects words that are similar in pronunciation and meaning between Arabic and other languages, this process helps to accelerate learning in an easy, effective and enjoyable way. Thus, the learner will find it easier for him to remember words with pronunciation similar to his mother tongue. This research aims to use the computer and information technology to accelerate the learning and teaching of the Arabic language to non-native speakers (the English language is a model in this research) by focusing on the similar words (cognates) between Arabic and English, through the program, the learner will be able to create and add lessons, then the cognates will be extracted automatically by computer. The analytical approach will be used to study the methods and algorithms used to find similarities between different languages, and then developing a new algorithm to find similarities between Arabic and English languages, and this method is considered one of the Natural Language Processing techniques.

Keywords: Cognates words, algorithm for finding Cognates words, Computer-Assisted Language learning.