

المعالجة الآلية للغة العربية المنطوقة

- واقع وآفاق -

Automatic processing of spoken Arabic

Current Status and Prospects

د.دفتيحة بوشان *

جامعة البويرة، كلية الآداب واللغات، الجزائر، f.bouchan@univ-bouira.dz

تاريخ الاستلام: 2026/02/07؛ تاريخ القبول: 2026/04/16؛ تاريخ النشر: 2026/01/30

ملخص:

تهدف هذه الدراسة إلى تسليط الضوء على واقع المعالجة الآلية للغة العربية المنطوقة، وهذا من خلال عرض أبرز التقنيات الصوتية الحديثة في مجال معالجة اللغات، والتطبيقات المتعلقة بها، ثم الاطلاع على حظ اللغة العربية من تلك الإنجازات، وأهم تحديات معالجتها آليا.

وقد توصلت الدراسة إلى أنّ المعالجة الآلية للغة العربية المنطوقة، قد قطعت أشواطاً هامة، حيث تمكنت من تطوير عدّة برامج وتطبيقات متعلّقة بتقنيات المعالجة الصوتية، كالمساعد الصوتي، وروبوتات الدردشة الصوتية، وتحويل الكلام إلى نص، والتعرّف على المتحدث، إلّا أنّ التحدي لا يزال مستمرا، وتطمح البحوث الحالية إلى تحقيق نتائج أفضل، إذا ما وقّرت لها الظروف الملائمة والإمكانات المادية اللازمة.

كلمات مفتاحية: معالجة اللغة المنطوقة؛ التقنيات اللغوية الصوتية؛ تطبيقات معالجة الكلام؛ تحديات اللغة العربية.

Abstract:

This study aims to shed light on the current state of spoken Arabic natural language processing (SLP). It presents the most prominent

modern acoustic technologies in the field and their related applications, and then explores the extent to which Arabic has benefited from these advancements and the most important challenges of its automatic implementation

The study concluded that the automatic processing of spoken Arabic has made significant strides, with the development of several programs and applications related to speech processing technologies, such as a voice assistant, Voicebots, speech-to-text conversion, and speaker recognition, but the challenge remains, and current research aims to achieve better results, if the appropriate conditions and necessary financial resources are provided.

Keywords: spoken language processing; speech technologies; speech processing applications; challenges of the Arabic language.

مقدمة:

يشهد العالم اليوم تطوراً تكنولوجياً مذهلاً مسّ مختلف المجالات، ومن مظاهر هذا التطور بروز تقنيات في غاية الأهمية، تمخضت عنها مجموعة من التطبيقات التي تسهل على الأفراد والجماعات نشاطهم، وتوفّر لهم الوقت والجهد.

ولأنّ مستقبل الشعوب مرهون بسلطة المعرفة، وقوامها التكنولوجية الرقمية، بات من الضروري مساندة هذا التطور، واستغلال ما يمكن أن تقدمه الآلة التكنولوجية للإنسان، عن طريق خلق حوار جاد بينه وبينها، ولا يمكن أن تقدّم الآلة المساعدة المطلوبة، ما لم تفهم لغة المحاور، ولا يمكن أن يتحقّق هذا إلّا ضمن ما يسمى بالمعالجة الآلية للغة الطبيعية.

لقد تمكّنت الأمم المنتجة للمعرفة والتكنولوجية، من إحراز تقدم كبير في مجال المعالجة الآلية للغتها، كما هو الحال مع اللغة الإنجليزية، في مقابل ذلك تقف الأمم المستقبلية للمعرفة دون إنتاجها، ومنها اللغة العربية، أمام مجموعة من التحديات، تفرضها أنظمة معدة في معظم الأحيان، لمعالجة لغة مخصصة غالباً ما تكون اللغة الإنجليزية، مما يستوجب الأمر تطوير هذه الأنظمة الأصلية، لتتماشى مع خصائص اللغة

التي يراد معالجتها.

إنّنا في هذه الدراسة سنتوقف عند المعالجة الآلية للغة العربية، المنطوقة منها دون المكتوبة، لمعرفة إلى أين وصلت الأبحاث في هذا المجال، وما الذي أنجز وما الذي ينتظر إنجازه.

للإجابة عن هذه التساؤلات وغيرها، اعتمدنا على خطة مبنية على المحاور الآتية:

- مفهوم معالجة اللغة الطبيعية
- التقنيات اللغوية الصوتية
- برامج وتطبيقات داعمة للغة العربية
- تحديات معالجة اللغة العربية المنطوقة

1. مفهوم معالجة اللغة الطبيعية (- Natural Language Processing)

:NLP

هي مجال فرعي من اللغويات، وعلوم الحاسوب، وهندسة المعلومات، والذكاء الاصطناعي، وتُعنى بالتفاعلات بين الحواسيب واللغات البشرية، وخاصةً كيفية برمجة الحواسيب لمعالجة وتحليل كميات كبيرة من بيانات اللغة، المكتوبة منها والمنطوقة⁽¹⁾.

ولأنّ البيانات اللغوية منها ما هو منطوق، ومنها ما هو مكتوب، فإنّ المعالجة اللغوية تتمّ ضمن قسمين هما⁽²⁾ :

- معالجة النصوص المكتوبة: تتم المعالجة في هذا القسم، باستخدام المعرفة المعجمية والنحوية والدلالية للغة، فضلاً عن أي معلومات مطلوبة في العالم الحقيقي.

- معالجة اللغة المنطوقة: تتمّ المعالجة في هذا القسم، باستخدام جميع المعلومات المطلوبة في القسم الأول، بالإضافة إلى المعرفة بعلم الأصوات، ومعلومات كافية للتعامل مع

(1) - Ankur A. Patel & Ajay Uppili Arasanipalai, Applied Natural Language Processing in the Enterprise, Teaching Machines to Read, Write & Understand, O'Reilly, 1st ed, United States of America, 2021, p4.

(2)-Patrick Doyle, Natural Language, From <https://users.cs.duke.edu/~brd/Teaching/Previous/AI/Lectures/Summaries/natlang.html> . Retrieved: 31/07/2025.

المزيد من الغموض الذي ينشأ في الكلام.

تهدف معالجة اللغة الطبيعية، إلى تطوير تطبيقات جديدة تسهّل التفاعل بين اللغة البشرية والحواسيب .

كما تهدف، من جهة أخرى، إلى نمذجة الآليات المعرفية الكامنة وراء إنتاج وفهم اللغة البشرية.⁽¹⁾

ومع التطور التكنولوجي الكبير، ظهر في المعالجة الآلية للغات ما يسمى بـ التّعلم العميق (Deep Learning)، الذي يُعد شكلاً متقدماً من التّعلّم الآلي، حيث يستند إلى قدرات الذكاء الاصطناعي، والهدف منه هو تقليد وظائف الدماغ عبر استيعاب كميات ضخمة من البيانات ومحاولة التعلّم منها، غير أنّه لا يزال هذا النوع من التعلّم إلى يومنا، غير قادر على محاكاة القدرات العقلية البشرية بصفة كاملة.⁽²⁾

2. التقنيات اللّغوية الصوتية:

1.2 التعرف على الكلام (-ASR Automatic Speech Recognition):

1.1.2 تعريفها:

هي تقنية تعد من أهم التقنيات اللغوية الصوتية، وأساس العديد من التطبيقات الصوتية، وأساس بعض التقنيات الأخرى كتقنية تحويل الكلام إلى نص.

يعبّر اسم هذه التقنية عن وظيفتها، وهي التعرف على الكلام البشري وفهمه، لقد كان استخدام اللغة المنطوقة للوصول إلى أجهزة الكمبيوتر هدفاً بحثياً مهماً منذ ستينيات القرن العشرين، ركزت الأبحاث في بداية الأمر على التعرف على الكلمات المنفصلة، قبل أن تتطور وتعالج الكلام المتّصل.⁽³⁾

ويبدو أنّ الاهتمام بالتّعرف الآلي على الكلام في اللغة العربية، كان خلال الفترة الممتدة

(1) -Iman Oraif, Natural Language Processing (NLP) and EFL Learning: A Case Study Based on Deep Learning, Journal of Language Teaching and Research, Vol. 15, No. 1, January 2024, p202.

(2) - Ibid, p202-203.

(3) - Patrick Doyle, Natural Language.

من 2004 إلى 2010، وكان هذا متزامنا مع البرامج التي كانت تمويلها وكالة مشروعات الأبحاث الدفاعية المتقدمة (DARPA)، التي عملت على تحسين نسبة التعرف الخاصة بأنواع مختلفة من الكلام المتدرجة الصعوبة، كما اهتمت هذه الهيئة بتمويل أبحاث اللغة العربية، وأدى هذا الأمر إلى تطوير نظم التعرف على الكلام العربية بشكل كبير، حيث تمكن الباحثون من بناء أنظمة للتعرف على الأخبار، والمحدثات التلفزيونية، وكذلك بعض اللهجات المحلية، مثل العراقية والشامية.⁽¹⁾

2.1.2 نظامها:

يتكوّن نظام التعرف على الكلام من⁽²⁾:

- بيانات أكوستيكية (Acoustic data): تتكوّن هذه البيانات من ملفات الكلام والنّص المصاحب لها، وعادة ما تتضمن هذه الملفات، عدة آلاف من الساعات للحصول على الدّقة المطلوبة، ليتم بعدها بناء نماذج الفون.

- بيانات لغوية: تتكوّن هذه البيانات من نصوص، ويفضل أن تكون نصوصا قريبة من تلك التي سيتم التعرف عليها لاحقا، فللتعرف على الأخبار على سبيل المثال، يفضل استخدام نصوص مأخوذة من الأخبار كالصحف مثلا. وتستخدم البيانات اللّغوية لبناء النماذج اللغوية.

وعليه فإنّ جوهر جميع أنظمة التعرف على الكلام، يتألف من مجموعة من النماذج الإحصائية، التي تمثل الأصوات المختلفة للغة المراد التعرف عليها، مع العلم أنّ نماذج ماركوف الخفية.

(Hidden Markov Models-HMM)، تُعدّ جوهر أغلب أنظمة التعرف على الكلام الحديثة، حيث توفر إطارا ملائما لبناء مثل هذه النماذج.⁽³⁾

(1) - محمّد عفيفي، معالجة النّص العربي المنطوق، فصل من: تطبيقات أساسية في المعالجة الآلية للغة العربية، الرياض، مجمع الملك سلمان العالمي للغة العربية، 1445هـ، ص 61-71.

(2) - المرجع نفسه، ص 70-71.

(3) - Mark Gales1 & Steve Young, The Application of Hidden Markov Models in Speech Recognition, Foundations and TrendsR in Signal Processing, Vol: 1, No3, 2007, p197.

- محرك بحث: بعد بناء النماذج الأكوستيكية واللغوية، يتم دمجها مع قاموس النطق لبناء فضاء البحث، وعند التعرف على الكلام، تستخدم أحد محركات البحث للبحث عن أفضل سلسلة من الكلمات.

3.1.2 مراحل المعالجة⁽¹⁾:

أ. ما قبل المعالجة:

- الالتقاط: يتم فيها التقاط الموجة الصوتية، لتحوّل إلى شكل قابل للتخزين.
- الترشيح: يميّز فيها صوت المتحدث، ليوضع ضمن قاعدة بيانات صوتية خاصة، وذلك وفقا لخصائصه الفيزيائية والنطقية.
- التقطيع: يتم فيها تجزئ العينة موضوع الدراسة.
- ب. المعالجة: يتم في هذه المرحلة، تحويل المعطيات المحصل عليها في المرحلة السابقة، إلى بيانات معالجة وفق لغة يفهمها الحاسوب لغة (0-1) .

ت. ما بعد المعالجة:

- التصنيف: تفرز في هذه المرحلة الأصوات اللغوية عن غيرها.
- التعرف: يتعرّف الحاسوب على الأصوات، عن طريق قيم الطيف الخاصة بكل موجة، والتي تمّ إدخالها ضمن قاعدة البيانات.
- الحفظ: آخر ما يقوم به الحاسوب هو حفظ المعلومات – الأصوات المعالجة- وتخزينها ضمن ملف لإخراجها وقت الحاجة.
- وعلى الرغم من التطور المستمر الذي تشهده تقنية التعرف على الكلام، إلا أنّ دقة نتائجها مرتبطة بعوامل مختلفة منها:⁽²⁾

(1) - يمينة زكري، التطبيقات الآلية لمعالجة الصوت- خطوة واقعية لحل إشكالية التواصل في ظل اللسانيات الحاسوبية، المقري للدراسات اللغوية والنظرية والتطبيقية، مج 03، العدد 02، 2020، ص69.

(2) - ينظر :

- Daniel Jurafsky & James H. Martin, Speech and Language Processing- An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models, 3rd ed draft, 2025, pp232-233, From .

- حجم الكلمات: كلما قلَّ عدد الكلمات التي يتضمنها الكلام، زادت دقة تقنية التعرف على الكلام.

- الجهة التي يوجه لها الكلام: يسهل التعرف على الكلام الموجه للآلات -كما في حالة التحدّث إلى نظام حوار-، مقارنةً بحديث البشر مع بعضهم البعض. إنّ الكلام الموجه للآلات عادة ما يكون مبسطاً وواضحاً وبطيئاً، على عكس الكلام الذي يكون موجهاً للبشر.

- الضوضاء: يسهل التعرف على الكلام، إذا تمّ تسجيله في غرفة هادئة، باستخدام ميكروفونات مثبتة على الرأس، مقارنةً بتسجيله بميكروفونات بعيدة، وفي مكان صاخب.

- اختلاف اللهجة أو خصائص فئة المتحدث: يسهل التعرف على الكلام إذا كان المتحدث يتحدث بنفس اللهجة التي تم تدريب النظام عليها، وينتهي إلى نفس الفئة التي تدرب عليها النظام (كفئة البالغين مقارنة بفئة الأطفال)، مع أخذ بعين الاعتبار أنّ اختلاف النطق قد نجده أيضاً عند المتحدث الواحد، الذي لا ينطق الكلمة نفسها دائماً بنفس الطريقة، وأحياناً تُسقط منه مقاطع صوتية عند حدود الكلمات. مع العلم أنّ تحديد حدود كلمات يعد إحدى المشكلات الصعبة في فهم الكلام.

بالإضافة إلى ما سبق، فإنّ الكلام قد لا يعبر دائماً عن المعنى الدقيق الذي يقصده المتحدث، وإنّما يفهم المعنى بمراعاة سياق الحال أيضاً، وهذا ما تعجز الآلة عن فهمه، ويتفوق عليها الإنسان في هذا الأمر.

4.1.2 أهميتها:

لقد تطورت تقنية التعرف التلقائي على الكلام، وأصبحت تستخدم لتأدية مهام عملية مختلفة منها:

- التواصل مع أجهزة المنزل الذكية أو المساعدين الشخصيين.
- تستخدم في تطبيقات الاتصالات الهاتفية نحو توجيه المكالمات والحوارات.
- تعتمد عليها تقنية النسخ الآلي الحديثة، في التعرف على الكلام قبل نسخه، أو

ترجمته.

- التواصل والتفاعل مع أجهزة الكمبيوتر، خاصة بالنسبة لذوي الإعاقة، ممن يعانون من صعوبات أو عجز في الكتابة أو السمع.

2.2 تحويل الكلام إلى نص (speech-to-text):

1.2.2 تعريفها:

تحويل الكلام إلى نص هي تقنية تُحوّل اللغة المنطوقة إلى كلمات مكتوبة، حيث تستخدم فيها أنظمة التعرف التلقائي على الكلام (ASR) خوارزميات معقدة، وشبكات عصبية لتحليل الإشارات الصوتية، وتحديد أنماط الكلام، وتحويلها إلى كلمات مكتوبة. تتعلم هذه الأنظمة من مجموعات بيانات ضخمة، مما يرفع من دقتها تدريجياً.

لم يقتصر استخدام هذه التقنية للتعامل مع محادثات الأفراد فقط، حيث تمكنت العديد من أنظمة هذه التقنية من التعامل مع عدة متحدثين، والتمييز بينهم، ونسب النص إلى قائله، مما يجعلها مثالية لنسخ المناقشات الجماعية.⁽¹⁾

ومع استمرار تطور هذه التقنية، تستمر معدلات الأخطاء في الانخفاض. على سبيل المثال، كانت معدلات الخطأ في نسخ الكلام الهاتفي حوالي 50% عام 1995، وفي حدود 2007 ومع توفّر المزيد من البيانات، والتحسينات فقد انخفضت معدلات الخطأ إلى أقل بكثير من 20% ⁽²⁾. أما اليوم، عصر التكنولوجيا الدقيقة، وبفضل النماذج اللغوية الضخمة (LLMs) والمعالجة المتطورة، لا شك أنّ معدل الخطأ فيها قد انخفض كثيراً، خاصة في الظروف المثالية ومع اللغات المدعومة جيداً كالانجليزية.

2.2.2 أهميتها:

تتميز التقنيات الحديثة لتحويل الكلام إلى نص بميزات عديدة منها:⁽³⁾

(1) - ISID, Speech to Text Transcription of speech into editable and searchable text, From <https://isid.com/technology/speech-to-text/> Retrieved :05/11/2025.

(2)- Mark Gales1 & Steve Young , The Application of Hidden Markov Models, p286.

(3) - لمعلومات أوفى، ينظر:

-ISID, Speech to Text- seamless Transcription technology.
=

- فهرسة النصوص وتنظيمها ضمن قاعدة بيانات، مما ييسر على المستخدم الوصول إليها.

- إمكانيات البحث عن الكلمات، وتحديد المواقع الدقيقة داخل النصوص، وهذا في وقت وجيز.

- الترجمة إلى لغة ثانية أو إلى لغات مختلفة، أي تحويل الكلام إلى نص أو نصوص متعدد اللغات (Multi-language transcription)، ويعدّ هذا النوع من الترجمة ذا أهمية، خاصة في الفعاليات المباشرة، والمؤتمرات والمقابلات، ونسخ الأفلام ومقاطع الفيديو، إذ تُتيح هذه التقنية الفهم الفوري لجمهور متنوع يتحدث لغات مختلفة.

- تحرير سهل الاستخدام، مع إمكانية تحرير النصوص بتنسيقات مختلفة، منها: JSON و SRT و TEXT.

- إمكانية دمج القواميس المتخصصة، مما يرفع من دقة استخدام المصطلحات.

- مساعدة ذوي الاحتياجات الخاصة.

للإشارة فإنّ أنظمة المعالجة المتطورة، بإمكانها اليوم- عكس ما كانت عليه سابقا- القيام بأكثر من عملية كالتعرف على الكلام، وتحويل الكلام إلى نص، أي أنّها لا تفصل بين التقنيتين، وكمثال عن ذلك: نموذج ويسبر (Whisper) من شركة (OpenAI)، هو نموذج تعلّم آلي من بين مهامه، التعرف على الكلام، والنسخ الصوتي المتعدد اللغات، واكتشاف اللغة.

2.3 التعرف على المتحدث (Speaker Recognition):

1.2.3 تعريفها:

تستخدم تقنية التعرف على المتحدث – كما تشير تسميتها- للتعرف على المتحدث، وذلك بالاعتماد على خصائص صوت المُتحدّث، التي هي بيانات بيومترية فريدة، تعتبر

<https://isid.com/technology/speech-to-text/>

- ISID, Translation technology : Multi-language transcription , From <https://isid.com/technology/translate/> Retrieved :05/11/2025

معرفات حيوية، كدرجة الصوت، ونبرته، وشكل المسالك الصوتية، وديناميكيات الكلام...

لهذه التقنية شكلان رئيسان: (1)

- شكل مرتبط بنص : وفيه يتم التحقق من هوية المتحدث، من خلال مقارنته بنص سابق له، تمّ فيه تدريب النظام على التعرف عليه، إلا أنّ هذه الطريقة تتطلب أن يلقي الشخص خطابا بلهجة تطابق النصّ المعد سلفا، وإلا لن تنجح عملية التعرف.

- شكل غير مرتبط بنص: وهو أكثر مرونة من الأوّل، حيث يتمّ فيه التّحقق من هوية المتحدث، من خلال حديثه العادي، وبأي لهجة كانت.

2.2.3 نظامها:

يعمل نظام هذه التقنية، على تحويل المدخلات الصوتية الخاصة بالمتحدّث إلى توليفة فريدة، لتُحفظ - بعد إزالة الضوضاء المحيطة بها- في قاعدة البيانات، على أساس أنّها تمثّل توقيعا صوتيا فريدا للمتحدث، وتجرى لاحقا مقارنة التسجيلات الجديدة، بالتوليفة الأصلية المحفوظة في قاعدة البيانات.

للعلم فإنّ هذه التقنية - على عكس التقنيتين السابقتين - تُستخدم مع مختلف اللّغات واللهجات. (2)

3.2.3 أهميتها:

تكمن أهمية تقنية التعرف على المتحدث في كونها سريعة ودقيقة في التعرف على هوية الأشخاص، ولأنّها تستخدم بيانات بيومترية فريدة خاصة بكل شخص يصعب تقليدها حتى الآن، فهي أيضا الأضمن في عملية التوثيق، مقارنة بالطرق التقليدية الأخرى، التي أصبحت عرضة لهجمات سيبرانية متكررة، لكن من جهة أخرى قد تستغل البيانات الصوتية الخاصة بالأشخاص، لإلحاق الضرر بهم، نحو ما هو واقع في الحروب العصرية، حيث يتمّ تتبّع تحركات الأفراد، بتتبع اتصالاتهم الصوتية التي تكشف عنها هذه التقنية.

(1) - ينظر: ماركو مسعد، "بصمة الصوت" علامة فارقة تجارية وعسكرية، المجلة، على الرابط:

https://www.majalla.com/node/309761/%D8%B9 تاريخ الاطلاع: 2025/11/25.

(2) - ISID ,Speaker ID tech: advanced voice recognition solutions. From <https://isid.com/technology/speaker-id/> Retrieved :05/11/2025

تستخدم تقنية التعرف على المتحدث في الدول المتقدمة، على نطاق واسع في العديد من المجالات، كالمصارف، وشركات الطيران الكبرى، والقطاع الصحي، والأمني، والعسكري، كما أنها تستخدم في الأجهزة المنزلية الذكية، والهواتف وغير ذلك.

3. برامج وتطبيقات داعمة للغة العربية:

بحكم أنّ اللغة وسيلة للتفكير، والتعبير عن الأفكار، ونقل الخبرات إلى الآخرين، والتواصل، والتعليم، وقضاء مختلف الحاجات الفردية والجماعية، فلا غرابة أن تسعى برامج وتطبيقات المعالجة اللغوية، نحو تحقيق هذه الغايات المختلفة، باختلاف نوعها.

1.3 ما له علاقة بالتعرف على الكلام:

يتمتع نظام التعرف على الكلام بالعديد من البرامج والتطبيقات، وضعت من أجل تحقيق أغراض مختلفة منها: التحكم والتوجيه، والبحث في المستندات الصوتية، والحوارات المنطوقة التفاعلية...

من بين أهم الإنجازات، في هذا المجال، التي تدعم اللغة العربية نذكر:

1.1.3 المساعدات الصوتية (Voice Assistants) منها:

- المساعد الصوتي من قوقل (Google Assistant): يرتبط هذا المساعد بأجهزة قوقل وخدماته،

وبالإضافة إلى المساعد الصوتي، يملك قوقل تطبيقات مختلفة متعلقة بالتعرف على الكلام منها:

- البحث الصوتي (Voice Search): سواء عن طريق جهاز الحاسوب، أو الهاتف الذكية، ويعمل البحث الصوتي أيضاً في "خرائط جوجل" (Google maps) لإعطاء نتائج حول الأمكنة.

- التحكم بالصوت (Voice Access): من بين أهم وظائف هذا التطبيق هي التحكم في شاشة الهاتف بالصوت فقط، فهو بديل عن استخدام الأصابع عند استعمال الهاتف.

- أليكسا (Alexa) من شركة أمازون: هي مساعد صوتي ذكي، تتوفر Alexa عبر مجموعة متنوعة من الأجهزة الذكية، توصف بأنها مديرة المنزل الذكي والتسوق، لأنّ من

أهم وظائفها التحكم في الأجهزة المنزلية الذكية مثل الإضاءة، التلفاز، المكيفات، الأقفال الذكية والكاميرات، ولأنها من شركة أمازون، فهي مبرمجة أيضا، لتقديم خدمات التسوق المختلفة، بالإضافة إلى ذلك تجيب أليكسا عن الاستفسارات التي تطرح عليها كتلك المتعلقة بأحوال الطقس وأخبار الرياضة، وغيرها.

● **سيربي من شركة أبل:** هو المساعد الذكي الخاص بشركة أبل، يأتي محملا على الأجهزة الخاصة بهذه الشركة. تكمن وظيفته في تنفيذ الأوامر الصوتية، والتحكم في الأجهزة المنزلية الذكية. والإجابة على الأسئلة، يقوم نظامه بتحديد الرد على المطلوب إما ⁽¹⁾.

- البحث في ذخيرة الهاتف أو الأنترنت لإيجاد مقاربة للمطلوب.
- تنفيذ أوامر محددة كالإتصال بشخص معلوماته متاحة على الهاتف.

2.1.3 روبوتات الدردشة الصوتية (Voicebots):

● **شات جي بي تي (ChatGPT) من شركة أوبن آي (OpenAI):** يعد من أشهر روبوتات الدردشة، وفيه نوعان من الدردشة: الدردشة الصوتية (Voice Chat)، وهي التي تهمنا هنا، بالإضافة إلى الدردشة النصية (Text Chat)، يدعم هذا الروبوت اللغة العربية وبعض لهجاتها، خاصة مع إطلاق نماذج جديدة، مثل (GPT-4o).

● **جيميني (Gemini) من شركة قوغل:** يعد المنافس القوي لشات جي بي تي، يمكن الوصول إليه عبر الويب، كما أنه متوفر كتطبيق مستقل خاص بالهواتف، وهو أيضا مدمج مع خدمات جوجل الأخرى مثل (Drive) و (Gmail).

● **كوبيولوت (Copilot) من شركة مايكروسوفت:** يعتمد على تقنيات (OpenAI)، متوفر على (الواب) كما أنه يأتي مدمجا في (الويندوز) في بعض الحواسيب والهواتف الذكية.

3.1.3 تطبيقات تعليمية مختلفة: يمكن الاستفادة من تقنية التعرف على الكلام لإنشاء تطبيقات خاصة بمجال التعليم وتحديدًا تعليم القراءة والنطق.

(1) - منصور بن محمد الغامدي، الصوتيات الحاسوبية، فصل من كتاب: مدخل إلى اللسانيات الحاسوبية، مجمع الملك سلمان العالمي للغة العربية، الرياض، 1445 هـ، ص28.

من تطبيقات تعليم النطق، على سبيل المثال، التي تمثل درجة عالية من التحدي، تلك المتعلقة بتعليم قراءة وتجويد القرآن الكريم، حيث تحتاج بعض قواعد النطق في علم التجويد، إلى درجة عالية من الحساسية من أجل الحكم بصحة تطبيق القاعدة التجويدية، مما يتطلب استخدام معالج آلي للتعرف على الكلام ذي درجة عالية جداً من الدقة. وقد أظهرت نتائج إحدى الاختبارات التي أجريت بغرض تقييم أداء برنامج تعليم التجويد، في الحكم على قراءة متعلم عادي، مقارنة بأداء معلّم بشري، أنّ نسبة الاختلاف بينهما في الحكم هي حوالي 4٪ فقط، وهي نفس نسبة اختلاف أي من الشيوخ مع إجماع باقي الشيوخ.

وفي اختبار آخر لقياس أثر برنامج تعليم التجويد على تعلم الفرد المبتدئ لبعض أحكام التجويد، أظهرت النتائج ارتفاع منحنى التعلّم لمجموعة المستخدمين بنسب كبيرة، حيث بدأ بنسب تقترب من 40٪ وارتفع إلى ما يزيد على 70٪ خلال بضع ساعات من التدريب باستخدام البرنامج، وهذا دليل على فعالية التقنية وكفاءتها⁽¹⁾.

2.3 ما له علاقة بتحويل الكلام إلى نص:

تُستخدم تطبيقات تقنية تحويل الصوت إلى نص مكتوب، لتحقيق أغراض مختلفة، كنسخ الاجتماعات، وإنشاء ترجمات للفيديوهات، ومساعدة ذوي الاحتياجات الخاصة، وتمكين الأوامر الصوتية في الأجهزة الذكية. وتحسين خدمة العملاء من خلال أنظمة الرد الصوتي التفاعلي. إنّها تطبيقات في غاية الأهمية، لأنّها تختصر الوقت والجهد وتمكن المستخدمين، ومنهم ذوي الاحتياجات الخاصة، من الاستفادة من خدمات مختلفة.

وإذا كانت بداية هذه التطبيقات مع شركة قوقل، فإنّه اليوم هناك العديد من الشركات والمؤسسات المهتمة بتطوير هذا النوع من التطبيقات، وما نعرضه فيما يأتي ما هو إلاّ عيّنة صغيرة ممّا هو متاح للمستخدم.

- نوتا (Notta) من شركة (Notta AI): (2) نوتا هو مُساعد ذكي، يعتبر من

(1) - لمعلومات أوفى، ينظر: شريف مهدي عبده، تطبيقات معالجة اللغة العربية في مجال التعليم، فصل من : تطبيقات أساسية في المعالجة الآلية للغة العربية، مجمع الملك سلمان العالمي للغة العربية، الرياض، 1445 هـ، ص 107 - 108.

(2) - Notta.ai, From <https://www.notta.ai/en> Retrieved :06/02/2026.

- الأدوات الرائدة التي تعتمد على الذكاء الاصطناعي للقيام بعدة مهام، منها:
- التفرغ الصوتي: حيث يقوم بتحويل الكلام المباشر أو الملفات المسجلة إلى نصوص مكتوبة.
 - التلخيص: كتلخيص الاجتماعات والمقابلات الطويلة.
 - ترجمة الكلام بعدة لغات.
- يمكن أن تستفيد من تطبيق نوتا شرائح مختلفة من المجتمع، نحو:
- الطلاب الذين يريدون نسخ المحاضرات والمقابلات، أو ترجمتها.
 - مندوبو المبيعات، والمستشارون الذين يديرون الاجتماعات أو المفاوضات.
 - العاملون عن بعد.
 - الإعلاميون كالصحفيين والمدونين.
- يعمل نوتا على جهاز الكمبيوتر، كما يعمل على الهاتف الذكي والجهاز اللوحي.
- لايف ترانسكرايب (Live transcribe) من شركة قوقل: يقوم (Transcribe) بتحويل التسجيلات الصوتية أو الكلام الفعلي إلى نص، يمكن استخدام هذا التطبيق مع الهاتف الذكي أو الكمبيوتر.
 - ترانسكربتور (Transkriptor):⁽¹⁾ هو تطبيق يحمل اسم الشركة التي وضعته، يقوم Transkriptor بتحويل الملفات الصوتية إلى نصوص، وله عدة مزايا منها:
 - نسخ مخصص لفئات مختلفة، كالفرق القانونية، ومقدمي الرعاية الصحية، وعلماء النفس، والمستشارين، ومديري تكنولوجيا المعلومات، والمحترفين الإعلاميين.
 - تلخيص المحتوى وتدقيق اللغات.
 - التعامل مع عدة متحدثين باستخدام تقنية تمييز المتحدثين، مما يسهل متابعة المحادثات في الاجتماعات وغيرها.
 - إمكانية الترجمة إلى لغات مختلفة.

(1) - ترانسكربتور، تحويل الصوت إلى نص، على الرابط: <https://transkriptor.com/ar> تاريخ الاطلاع:

2026/01/20.

- إمكانية تحليل مشاعر المتحدثين.

• جي بورد (Gboard) من شركة قوقل: صمم (Gboard) خصيصا للأجهزة المحمولة (Android) و (iOS) وأجهزة (Android TV).

يستخدم هذا التطبيق خاصة في المهام اليومية والسريعة، ككتابة رسائل "واتساب" بالصوت، أو رسائل "البريد الإلكتروني"، بالإضافة إلى أنه يقوم بتفريغ المحاضرات والاجتماعات، وغير ذلك.

3.3 ما له علاقة بالتعرف على المتحدث:

من المنتجات المتاحة في أسواق البرمجيات ذات العلاقة بالتعرف على المتحدث:

Open Sesame, Nuance Vocal Password, Authentify, Voice Vault, iAM, Voice Print System Bio Validation, Voice Biometrics Group.

تستجيب هذه الأنظمة لصاحب الصوت المبرمج عليها فقط، وتستخدم تطبيقاتها لأغراض مختلفة كالدخول إلى الحاسبات الشخصية، أو الهواتف الذاتية، أو حسابات على الأنترنت، أو فتح أبواب البيت أو الغرف وغير ذلك.⁽¹⁾

وبالإضافة إلى الخدمات التي تقدمها هذه التطبيقات للأفراد، فإنّ هناك قطاعات مختلفة تستفيد من خدماتها، منها التحقيقات الجنائية، والمصارف والقطاع الصحي وشركات الطيران والتوظيف، كما استخدمت في الآونة الأخيرة بشكل كبير في العمليات العسكرية، خاصة في قطاع غزة، لتعقب قيادات "حماس" وتحديد أماكنهم قبل تصفيتهم.⁽²⁾

إنّ البيانات البيومترية التي تعتمد عليها تقنية التعرف على المتحدث، قد توظف في تطبيقات معينة لخدمة صاحبها، لكن قد توظف في تطبيقات أخرى لأغراض تعود بالضرر عليه.

مع الإشارة إلى أنّ نظام هذه التقنية مستقل عن أي لغة، وبالتالي يمكن للغة العربية

(1) - منصور بن محمد الغامدي، الصوتيات الحاسوبية، ص32.

(2) - لمعلومات أوفى عن استخدام هذه التقنية في المجال العسكري، ينظر: ماركو مسعد، "بصمة الصوت" علامة فارقة تجارية وعسكرية.

الاستفادة من مختلف تطبيقاتها.⁽¹⁾

4. تحديات معالجة اللغة العربية المنطوقة:

رغم كثرة التطبيقات التي تدعم اللغة العربية المنطوقة ولهجاتها، إلا أنه علينا أن نشير إلى أن فاعليتها لم تبلغ الدرجة التي بلغتها مع بعض اللغات كاللغة الإنجليزية، وهذا يعود إلى مجموعة من التحديات يمكن أن نصنفها إلى صنفين:

1.4 تحديات من داخل اللغة:

أ. غياب التشكيل في النصوص العربية: إن أغلب النصوص العربية تكتب بدون تشكيل، وهذا ما يشكل تحدياً لنظم التعرف على الكلام، حيث يصعب عليها التعرف على بعض الكلمات التي لا تكون مشكلة في قاعدة بياناتها.

ب. البناء الصرفي للغة العربية: تعتبر اللغة العربية لغة غنية صرفياً مقارنة باللغات الأخرى كالإنجليزية، وهذا ما خلق تحدياً أمام أنظمة معالجة اللغة العربية، في تغطية جميع الأشكال الممكنة للكلمات، حتى بقائمة مفردات كبيرة، ويمكننا أن نتصور مدى غنى اللغة العربية بالمفردات، إذا علمنا أن قاموساً مكوناً من 64 ألف كلمة يكفي لتغطية 99% من نصوص الأخبار في اللغة الإنجليزية، بينما نحتاج إلى عشرة أمثال هذا العدد أي حوالي 600 ألف كلمة للوصول إلى نفس النسبة في اللغة العربية.⁽²⁾

وقد أثبتت إحدى الدراسات أنه حتى باستخدام عدد كبير من الكلمات في أنظمة معالجة اللغة العربية، تبقى نسبة الكلمات الخارجة عن القائمة المضبوطة في النظام مرتفعة، مقارنة باللغة الإنجليزية على سبيل المثال، فباستخدام مفردات يبلغ حجمها 350 ألفاً في أنظمة معالجة اللغة العربية، بلغ معدل الكلمات خارج قاموس النظام 2.0%، وباستخدام 59 ألف كلمة فقط في أنظمة معالجة اللغة الإنجليزية، بلغ معدل الكلمات خارج قاموس النظام أقل من 2.0%.⁽³⁾

(1) - ISID, Speaker ID tech : advanced voice recognition solutions.

(2) - محمد عفيفي، معالجة النص العربي المنطوق، ص73.

(3) - لمعلومات أوفي ينظر:

-Mark Gales1 & Steve Young , The Application of Hidden Markov Models, in Speech Recognition, p282.

ت. خصوصية اللّغة العربية الصوتية: للغة العربية مجموعة من الخصائص الصوتية، كالإخفاء الشفوي والإقلاب والإبدال، والتي ينبغي أن تؤخذ بعين الاعتبار عند وضع أنظمة معالجة اللغة العربية المنطوقة. إنّ عدم مراعاة هذه القضية يؤثر سلباً على فعالية التطبيقات، وهذا ما وقع فعلاً خاصة مع التطبيقات الخاصة بنسخ الكلام المنطوق التي تعرّض الكثير منها إلى النقد نتيجة هذا الأمر⁽¹⁾.

ث. تنوّع اللهجات العامية: إنّ للغة العربية عدة لهجات مختلفة، ويتخطى هذا الاختلاف الكلمات إلى الفونات، ومما يزيد المشكلة تعقيداً أن اللهجات لا تكون مكتوبة، وبالتالي لا توجد نصوص كافية لبناء النماذج اللغوية. إنّ فهم كل لهجة يتطلب بيانات ضخمة وممثلة لخصوصياتها اللغوية، وهذا ما يصعب توحيد المعايير وبناء حلول شاملة، كما يزيد من تكلفة ووقت بناء النماذج.

2.4 تحديات من خارج اللغة⁽²⁾:

أ. محدودية المحتوى الرقمي العربي على شبكة الإنترنت: فبحسب إحصاءات منصة "دبليو ثري تكس" للتدريب وتوفير الاستشارات حول تطوير البرمجيات، فإنّ نسبة المحتوى العربي من إجمالي المحتوى العالمي على شبكة الإنترنت هو 0.6%، في مقابل 50.1% للغة الإنجليزية، وما من شك أنّ هذا الضعف في المحتوى ينعكس سلباً على معالجة اللغة العربية، وعلى التطبيقات الناجمة عنها.

ب. نقص الموارد اللّغوية المفتوحة المتاحة بالعربية: إنّ النقص على هذا المستوى، يطرح مشكلة بالنسبة لتكنولوجيا الذكاء الاصطناعي، عندما يتعلّق الأمر بمعالجة هذه اللّغة مقارنة بلغات الأخرى مثل الإنجليزية.

(1) - ينظر: أحمد علي علي لقم، تطبيقات هندسة اللغويات العربية واقع وآفاق، مجلة كلية اللغة العربية بإتاي البارود، المجلد 31، العدد 1، 2018، ص 1125-1173، جامعة الأزهر، مصر، ص 1164-1165.

(2) - ينظر :

- رنا الدقاق، اللسانيات الحاسوبية...البدايات والتطبيقات والتحديات، رابط الموقع: <https://trendsresearch.org> تاريخ الاطلاع: 2025/07/29.

- جامعة محمد بن زايد للذكاء الاصطناعي، الذكاء الاصطناعي واللغة العربية: فرصة لتعزيز الهوية الثقافية، تاريخ الاطلاع: 2025/07/13، على الرابط: <https://mbzuai.ac.ae>

ت. قلة مصادر البيانات الضخمة المتعلقة باللغة العربية: لا شك في أنّ قلة مصادر البيانات الضخمة (Big Data)، يعيق تحقيق أداء عالٍ لتطبيقات الذكاء الاصطناعي، القدرة على معالجة هذه اللغة في كل مستوياتها الثقافية والدلالية، سواء كانت مكتوبة أو مسموعة أو مرئية، الأمر الذي يستدعي ضرورة العمل على توفير بيانات متنوعة وشاملة، وذات جودة عالية .

ث. نقص الكوادر المؤهلة: يمثل نقص الكوادر في هذا النوع من العلوم تحديًا آخر، فتعليم علوم الحاسوب في معظم نُظم التعليم المدرسية في الدول العربية، يقتصر على مبادئ أوليّة، لا ترقى إلى التعريف بعلوم البرمجة، وحتى إن وُجدت هذه الكوادر، فإنها تهاجر للالتحاق بالشركات العالمية الكبرى، كما أنّ معظم الجامعات العربية لم تستحدث بعد أقسامًا في كلياتها تتيح التخصص في الدراسات اللغوية الحاسوبية.

ج. نقص الدعم المادي: إنّ البحث العلمي في معظم الدول العربية، لا يحظى بالدعم المادي الذي يحتاج إليه، ومعلوم أنّ التمويل في هذا المجال، يمثل البنية التحتية لعمليات التطوير والبحث، ونقصه يؤدي إلى غياب الجانب التطبيقي فيه، وبالتالي غياب الإنتاج الفعلي للبحوث.

خاتمة:

في ختام هذا البحث نخلص إلى جملة من النتائج نستعرضها في النقاط الآتية:

- لقد أحرزت المعالجة الآلية للغة العربية المنطوقة تقدماً ملفتاً، خاصّة في السنوات الأخيرة، يظهر هذا من خلال تطوير تطبيقات عديدة لها علاقة بتقنيات مختلفة، كالترعرع على الكلام، وتحويل الكلام إلى نص، والتعرّف على المتحدث، إلّا أنّه علينا أن نقرّ أنّ معالجة اللغة العربية، لم تبلغ بعد الدقّة المرجوة، مقارنة بما بلغته في لغات أخرى كاللغة الإنجليزية.

- إنّ التحدي القائم في معالجة اللغة العربية عموماً، نابع من اختلاف خصائص اللغة العربية عن خصائص اللغة المنتجة للمعرفة والتقنية، ومنها اللغة الإنجليزية على وجه الخصوص، وهذا يعني ضرورة تطوير البرامج الأصلية، لتتّمشى مع خصائص اللغة العربية، مما يتطلب الأمر بذل المزيد من الجهود وتوفير الإمكانيات المادية والمعنوية اللازمة للباحثين في هذا المجال.

- لا شك في أنّ التقدّم الذي أحرزته المعالجة الآلية للغة العربية المنطوقة، يفتح آفاقا واعدة لتطبيقات الذكاء الاصطناعي باللغة العربية، في مختلف المجالات، كالبحث العلمي، والترجمة الآلية عالية الجودة، والتعليم، وصناعة المحتوى الرقمي... لكن علينا أن نحترس من الجانب السلبي لهذه التطبيقات، الذي قد يستغله الأكثر علما في هذا المجال، ويذهب ضحيته المستهلك الذي يجهل خبايا هذه التطبيقات، لأنه لا يملك أسرار التكنولوجيا الحديثة، هذا من جهة، ومن جهة أخرى، لا توجد ضمانات ولا قوانين تحكم استخدام تطبيقات الذكاء الاصطناعي، وتحمي حقوق وخصوصية الأفراد والجماعات والدول أيضا، وقد شاهدنا في السنوات الأخيرة، كيف استغلت تطبيقات مختلفة في إدارة الحروب، وخلق النزاعات، وابتزاز الأشخاص، فتحوّلت بذلك التكنولوجيا إلى نقمة في أيدي من لا ضمير لهم ولا أخلاق.

وبناء على نتائج هذا البحث، فإننا نوصي بـ:

- توفير الدعم المادي والمعنوي، خاصة للجامعات العربية، من أجل تطوير البحوث المتعلقة بمعالجة اللغة العربية المنطوقة، وتأهيل كوادر متخصصة في هذا المجال، تدفع بهذه البحوث قُدماً نحو تطوير تطبيقات ذكية، تكون قادرة على القيام بمهام مختلفة ودقيقة، وتواجه مختلف التحديات.

- التخطيط اللغوي لمواجهة تحديات معالجة اللغة العربية في ضوء استراتيجية الذكاء الاصطناعي.

- الاستفادة من مختلف الخدمات التي يمكن أن تقدمها التطبيقات الصوتية، وأن يوسع استخدامها لتشمل مجالات أخرى ذات أهمية، كمدال التعليم، الذي لم ينل حظه الكافي من هذه التطبيقات.

- تدعيم المحتوى الرقمي باللغة العربية.

- توفير الموارد المفتوحة من أجل تطوير البحث المحوسب للغة العربية، وتشمل هذه الموارد مجموعات البيانات الضخمة، والموارد النصية المفتوحة، التي تسهّل على الباحثين الوصول إلى المعلومات الضرورية.

قائمة المراجع

باللغة العربية

أولاً: الكتب

1. شريف مهدي عبده، تطبيقات معالجة اللغة العربية في مجال التعليم، فصل من: تطبيقات أساسية في المعالجة الآلية للغة العربية، مجمع الملك سلمان العالمي للغة العربية، الرياض، 1445 هـ
2. محمد عفيفي، معالجة النص العربي المنطوق، فصل من: تطبيقات أساسية في المعالجة الآلية للغة العربية، الرياض، مجمع الملك سلمان العالمي للغة العربية، 1445 هـ
3. منصور بن محمد الغامدي، الصوتيات الحاسوبية، فصل من كتاب: مدخل إلى اللسانيات الحاسوبية، مجمع الملك سلمان العالمي للغة العربية، الرياض، 1445 هـ

ثانياً: المقالات العلمية

1. أحمد علي علي لقم، تطبيقات هندسة اللغويات العربية واقع وأفاق، مجلة كلية اللغة العربية بإتاي البارود، المجلد 31، العدد 1، 2018، جامعة الأزهر، مصر.
2. يمينه زكري، التطبيقات الآلية لمعالجة الصوت- خطوة واقعية لحل إشكالية التواصل في ظل اللسانيات الحاسوبية، المقري للدراسات اللغوية والنظرية والتطبيقية، مج 03، العدد 02، 2020.

رابعاً: المواقع الإلكترونية والمنصات الرقمية

1. <https://isid.com/technology/speaker-id/> Retrieved :05/11/2025
2. ISID ,Speaker ID tech: advanced voice recognition solutions. From
3. ISID, Speech to Text- seamless Transcription technology. <https://isid.com/technology/speech-to-text/>
4. ISID, Speech to Text Transcription of speech into editable and searchable text, From <https://isid.com/technology/speech-to-text/> Retrieved :05/11/2025.
5. ISID, Speaker ID tech : advanced voice recognition solutions.
6. ISID, Translation technology : Multi-language transcription , From <https://isid.com/technology/translate/> Retrieved :05/11/2025
7. Notta.ai, From <https://www.notta.ai/en> Retrieved :06/02/2026.

8. Patrick Doyle ,Natural Langage, From <https://users.cs.duke.edu/~brd/Teaching/Previous/AI/Lectures/Summaries/natlang.html> . Retrieved : 31/07/2025.

9. ترانسكربتور، تحويل الصوت إلى نص، على الرابط <https://transkriptor.com/ar> تاريخ الاطلاع: 2026/01/20.

10. جامعة محمد بن زايد للذكاء الاصطناعي، الذكاء الاصطناعي واللغة العربية: فرصة لتعزيز الهوية الثقافية، تاريخ الاطلاع: 2025/07/13، على الرابط <https://mbzuai.ac.ae>

11. رنا الدقاق، اللسانيات الحاسوبية...البدايات والتطبيقات والتحديات، رابط الموقع : <https://trendsresearch.org> تاريخ الاطلاع: 2025/07/29.

12. ماركو مسعد، "بصمة الصوت" علامة فارقة تجارية وعسكرية، المجلة، على الرابط : <https://www.majalla.com/node/309761/%D8%B9>، تاريخ الاطلاع: 2025/11/25.

باللغة الأجنبية

Books

1. Ankur A. Patel & Ajay Uppili Arasanipalai , Applied Natural Language Processing in the Enterprise ,Teaching Machines to Read, Write & Understand, O'Reilly, 1st ed, United States of America, 2021

2. Daniel Jurafsky& James H. Martin, Speech and Language Processing- An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition with Language Models, 3rd ed draft, 2025, From Retrieved : 31/07/2025 <https://web.stanford.edu/~jurafsky/slp3/>

3. Mark Gales¹ & Steve Young, The Application of Hidden Markov Models in Speech Recognition, Foundations and TrendsR in Signal Processing, Vol :1, No3, 2007

4. Mark Gales¹ & Steve Young, The Application of Hidden Markov Models, in Speech Recognition.

Scientific articles

1. Iman Oraif, Natural Language Processing (NLP) and EFL Learning: A Case Study Based on Deep Learning, Journal of Language Teaching and Research, Vol. 15, No. 1, January 2024