الدعم الصوتي باللغة العربية في خدمات جوجل

فائق عويس وكريستينا الحايك

المقدّمة

تتمثل مهمة شركة Google "جوجل" في تنظيم المعلومات حول العالم وتسهيل الوصول إليها والإفادة منها عالميًا، وضمن هذه الرسالة تأتي جهود الشركة في دعم اللغة العربية، حيث توفر خدماتها وبرامجها معرّبة وبلغة سليمة وطبيعية تحترم المستخدم العربي وثقافته. ولا تقتصر هذه الجهود على توفير واجهة مستخدم باللغة العربية لبرنامج ما وحسب، بل تتضمن أيضًا توفير الخدمات والبرامج والحلول كمنصات للمستخدم العربي تتيح له إنشاء وتطوير محتوى خاص به. هذا المحتوى، والذي يشتمل على المرئي والمقروء والمسموع وغيره، يصب في إنشاء بيئة معرفية يستفيد منها كل المستخدمين.

تسعى شركة جوجل إلى تقديم المساعدة لجميع الناس في جميع جوانب حياتهم اليومية، انطلاقًا من تسهيل التواصل مع الآخرين في أنحاء العالم ووصولاً إلى تنظيم يوم عملهم ليتسنى لهم التركيز على الأمور المهمة فقط. وفي هذا السياق، وبالإضافة إلى دعم اللغة العربية نصيًا وفي واجهة الاستخدام لمعظم منتجات وخدمات جوجل، فقد عملت الشركة على دعم الإدخال والإخراج الصوتي باللغة العربية في عدد من هذه الخدمات والأدوات لتسهيل حياة المستخدم وتوفير تجربة حول معنى أن تكون حياتنا متصلة بالإنترنت وأن تصبح المعلومات في متناول أيدينا عندما نكون في أمسّ الحاجة إليها.

في هذه الورقة سنناقش التحديات التي واجهت دعم اللغة العربية وخاصة الدعم الصوتي والتعرف على الكلام وبلهجات عربية مختلفة والحلول التي وضعتها جوجل وأتاحت هذا الدعم في العديد من منتجاتها وخدماتها.

ما هي تكنولوجيا الكلام؟

قبل سنوات ليست بالبعيدة جدًّا، كان تواصل الإنسان مع الآلة أو الكمبيوتر ضرب جنون أو خيال، ولكن ما كان في الماضي نسج تصوّرٍ مستقبليٍّ لريشة كاتب أو حيلة تأثيرات خاصة لصانع أفلام سابق لعصره، أصبح اليوم حقيقة تتجلى على شكل كمبيوتر محمول أو هاتف جوال أو جهاز لوحي أو ساعة في اليد أو حتى نظارة.

وصلت هذه الأجهزة، وخاصة المتصلة بالإنترنت منها، إلى درجة من الذكاء يجعل منها أشبه بمساعد شخصي يعرف اهتماماتك واحتياجاتك وينظّم مهامك ويحفظ أفكارك ويقترح عليك الحلول من دون حتى أن تسأل أحيانًا، وفي حال سألت، تراها تفهم كلامك وتعطيك الإجابات سواء على شكل نص مكتوب أم على شكل كلام منطوق يشبه إلى حدٍّ كبير لغتك الشفهية.

وتسأل كيف لجهازٍ لا حياة ولا عقل إنسان ولا مشاعر له أن ينطق بنتائج بحث ما أو يقرأ لك مقالة أو يقول لك إرشادات الطريق، ويستوعب في الوقت عينه كلامك؟

الأمر ممكن بفضل تكنولوجيا الكلام وهي مجموعة تقنيات تم تصميمها لتقليد الصوت البشري والرد عليه، ويتم تطبيق هذه التكنولوجيا في استخدامات مختلفة منها مساعدة الأشخاص ذوي الاحتياجات الخاصة كالصم والبكم والمكفوفين، بالإضافة إلى التواصل مع الكمبيوتر من دون استخدام لوحة المفاتيح، كما أنّ هذه التقنيات تسهم في تحسين برامج الألعاب الإلكترونية والخدمات عبر الهاتف.

يتفرّع من هذه التكنولوجيا مجالات عديدة منها:

- التعرف على الكلام وهو يندرج ضمن علم اللغويات الحاسوبية الذي يجمع بين المعرفة والأبحاث في اللغويات وعلوم الكمبيوتر والهندسة الكهربائية لتطوير أساليب وتقنيات تمكّن التعرف على اللغة المحكية وترجمتها على شكل نص مكتوب من خلال الكمبيوتر. ويطلق عليها أيضًا اسم "التعرف الآلي على الكلام" أو "تعرّف الكمبيوتر على الكلام" أو "تحويل النص إلى كلام". بعض أنظمة التعرف على

الكلام تعتمد "التدريب" بحيث أنّه كلما أدخل متكّلم ما كلامًا في النظام، يحلل النظام صوت المتكلم ويستخدمه لتحسين التعرف على هذا الصوت ما ينتج عنه دقة أكبر في تحويل كلامه إلى نص.

- توليف الكلام وهو الإنتاج الإصطناعي للكلام البشري ويعتمد على نظام كمبيوتر يسمى كمبيوتر كلام أو محرّك توليف كلام يمكن تطبيقه في البرامج والأجهزة الإلكترونية. يعمل هذا النظام بشكل أساسي على تحويل النص المكتوب إلى كلام. وتهدف تقنيات توليف الكلام إلى ترجمة سلسلة من الرموز الصوتية وتحويل رمز لغوي ما من أجل توليد الكلام آليًّا باستخدام معلومات عن النبرة والإيقاع.
- التعرف على المتكلم هي طريقة التعرف على الشخص المتكلم من خلال الإشارات الصوتية في الكلام وهي تختلف عن طريقة التعرف على الكلام نفسه، بل على هوية المتكلم – من هو هذا الشخص. والتعرف على المتكلم يأتي من خلال الميزات الصوتية والتي تختلف من شخص إلى آخر.
- التحقق من المتكلم هي عملية التحقق من والمصادقة على هوية المتكلم من خلال فهم الكلام إذا كان لهذا الشخص المتكلم أم لا.
- ترميز الكلام وهي عملية تشفير وضغط البيانات الصوتية الرقمية التي تحتوي على كلام وتستخدم لمعالجة الإشارات الصوتية جنبا إلى جنب مع خوازميات معينة للخروج بنموذج صوتي في أجهزة الهاتف أو الحواسيب أو أجهزة نقل الصوت.

تكنولوجيا الكلام في منتجات جوجل

تطبّق جوجل تكنولوجيا الكلام على أشكالها المختلفة في عدد كبير من منتجاتها، وتسعى بشكل خاص إلى توفير هذا الدعم باللغة العربية. في ما يلي لمحة عامة عن هذه المنتجات التي تسمح للمستخدم التفاعل معها بالصوت.

أوّلاً: البحث الصوتي (في نظام تشغيل الهواتف Android ومتصفح Chrome)

هو استخدام الميكرفون في جهاز الكمبيوتر أو الهاتف الذكي لطلب بحث معين بدلاً من كتابته. خدمة البحث الصوتي باللغة العربية في منتجات جوجل متوفرة حاليًّا بثلاثة عشر لهجة عربية منها: السعودية، المصرية، الأردنية، الفلسطينية، اللبنانية، الإماراتية، الكويتية والقطرية، ويجدر بالذكر أن اللغة الفصحى مدعومة في جميع الإصدارات.

ومن الأمثلة على الأدوات التي تدعم البحث الصوتي تطبيق Google Now وهو المساعد الشخصي والذي يمكنه تزويد المستخدم بمعلومات تناسب احتياجاته، سواء من خلال تنبيهات بتجنب السير في بعض الطرق نظرًا لازدحام حركة المرور فيها، أو من خلال إخبار المستخدم بكل ما هو جديد عن حالة الطقس.

بطاقات المساعد الشخصي Google Now

بعد النقر على رمز الميكروفون، ينطق المستخدم العبارة التي يريد البحث عنها. في هذا المثال، يطلب المستخدم معلومات عن "الطقس في الرياض" فيقوم البرنامج بعرض بطاقة تتعلق بحالة الطقس مثل حالة الأمطار والرطوبة والرياح ودرجة الحرارة لعدة أيام.

اللهجات المختلفة للبحث الصوتي

ثانيًا: خرائط جوجل Google Maps

توفّر خدمة خرائط جوجل إلى المستخدمين المساعدة في الوصول إلى وجهاتهم المحلية المختلفة عن طريق إظهار خريطة المسار الدقيقة من نقطة الانطلاق إلى الوجهة النهائية. ويستطيع المستخدم إدخال وجهته بالصوت ومن دون الاستعانة بلوحة المفاتيح. وبالتالي يحصل على إرشادات التنقل مكتوبةً فوق المسار على الخريطة ومنطوقةً من الجهاز بحيث يكفي له الاستماع للإرشادات باللغة العربية الفصحى فيسهل عليه الوصول إلى العنوان المنشود من دون أن يتشتت انتباهه عن الطريق.

Al Bustan Rotana
1 س 25 د
1 س 25 د (١٥٠ كم)

البستان روتانا دبي
Prime Hospital
Millennium Airport Hotel Dubai
Jules Bar
Meriden Village Terrace
Al Bustan Rotana
US$ 190
51st St
Google
Al Bustan Rotana

المتابعة في خط مستقيم
٧٥٠ متر
Widheel St
شارع هزاع بن زايد
1 س 25 د
إعادة توسيط

ثالثًا: ترجمة جوجل

هي خدمة مقدمة من جوجل لترجمة جزء من نص أو صفحة إنترنت إلى لغة أخرى وتعتمد على أسلوب الترجمة الآلية الإحصائية، حيث تقوم بالبحث في عدد هائل يصل لمئات الملايين من المستندات المتوفرة على الإنترنت وذلك لتحديد أفضل ترجمة للنص المطلوب. و تقوم ترجمة جوجل بتخمينات ذكية لتحديد الترجمة المناسبة. وللحصول على هذه الكمية الضخمة من البيانات اللغوية، تستخدم Google العدد الكبير من النصوص الثنائية اللغة الموجودة على الإنترنت مثل وثائق الأمم المتحدة التي تتوفر في ست لغات بشكل متواز، بما فيها اللغتين العربية والإنكليزية. تتميز خدمة الترجمة بأنها تتيح للمستخدم اختيار وضع المحادثة حيث يمكنه نطق الكلمات أو الجمل التي يريد ترجمتها وبالتالي الاستماع إلى ترجمتها.

خدمة الترجمة على الكمبيوتر وميزات الدعم الصوتي

- استخدام رمز الميكروفون لإدخال (نطق) النص المراد ترجمته
- استخدام رمز "السماعة" للاستماع إلى النص الذي تم إدخاله
- ميزات أخرى مثل الكتابة الصوتية والكتابة اليدوية واستخدام لوحة المفاتيح أو كتابة الحروف العربية بحروف لاتينية
- إمكانية تحسين الترجمة واقتراح ترجمة أفضل

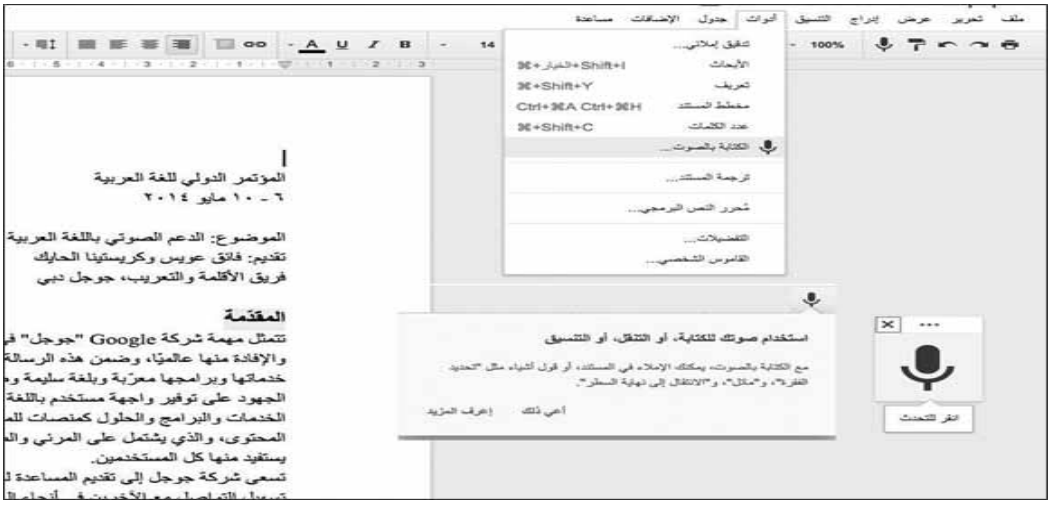
أما تطبيق الترجمة على الهواتف الذكية فقد أصبح بمثابة المترجم الشخصي والفوري لأي شخص يحمل هاتفًا ذكيًا. فبعد تحديد لغتين واستخدام "وضع المحادثة" بإمكان أي شخصين أن يتبادلا الحديث كل بلغته من خلال النطق والاستماع. ويمكن أيضًا الاستفادة من هذه الميزة في تعليم اللغة العربية لغير الناطقين بها. ومع تزايد المحتوى العربي على الإنترنت، أصبحت دقة وجودة الترجمة أفضل بكثير مما كانت عليه قبل سنوات.

تطبيق الترجمة للهواتف الذكية

- الإدخال والإخراج الصوتي
- الكتابة اليدوية أو بلوحة المفاتيح
- ترجمة الصور (من الإنجليزية والألمانية إلى العربية)
- المترجم الشخصي – وضع المحادثة

رابعًا: مستندات Google

هو برنامج يسمح للمستخدمين بإنشاء وتحرير الملفات والمستندات عبر الإنترنت بالتشارك مع مستخدمين آخرين في الوقت عينه، مع إمكانية إدخال النص من دون استخدام لوحة المفاتيح ولكن من خلال إملاء الكلمات شفهيًا ويحوّلها محرر المستندات إلى كلمات مكتوبة.



الكتابة بالصوت في مستندات Google

خامسًا: الكتب المسموعة

الكتب الصوتية أو المسموعة أصبحت تتوفر بشكل كبير على الإنترنت ولكن ما زالت نسبتها ضئيلة مقارنة بالكتب التي يمكن تصفحها وقراءتها. إن تكنولوجيا تحويل النص إلى كلام أصبحت تعطي دقة وجودة عالية في إخراج الصوت والذي سيوفر طريقة أخرى للتسجيل الصوتي التلقائي للكتب المتوفرة بصيغة نصية. وهذه الميزة ستوفر الكثير من الوقت والموارد، فبدلا من التسجيل التقليدي والمواهب والمعدات اللازمة للتسجيل واحتمالات الأخطاء اللغوية الناتجة عن القراءة التقليدية، فإن هذه الميزة تلقائية وهي نتيجة جهود وخبرات كبرى للشركات العالمية.

إنشاء أنظمة الإدخال والإخراج الصوتي

سوف نسلط الضوء في هذا القسم على العملية التي يتم من خلالها توفير دعم الإدخال والإخراج الصوتي في التطبيقات مع التركيز على تطوير هذا الدعم للغة العربية وما رافق هذا التطوير من تحديات.

بدايةً، نعرض بعض التفاصيل عن عملية التعرف على الكلام وهي عملية تعتمد على نظام يسمى بمحرك التعرف على الكلام الذي يأخذ إدخال دفق الكلام ويسجّله على شكل نص. يقوم النظام بتحليل دفق الكلام فيعزل المقاطع الصوتية ليحوّلها إلى سلسلة من الرموز الرقمية التي تمثّل الأصوات الملفوظة في الإشارة الصوتية.

بعد تحليل الكلام وتحويله إلى رموز رقمية، يبحث المحرك في ثلاث أنواع من قواعد البيانات هي: النموذج السماعي والقاموس، والنموذج اللغوي:

- نموذج السماعي يمثّل الأصوات السماعية الخاصة بلغة ما، ويمكن تدريبه للتعرف على الخصائص التي تميّز أنماط كلام المستخدم وبيئته السماعية.
- القاموس يضم عددًا كبيرًا من الكلمات الخاصة باللغة ويزوّد معلومات عن كيفية لفظ كل منها.
- نموذج اللغة يمثّل طرق دمج الكلمات وتجميعها في الجمل حسب الاستخدام الطبيعي لها في اللغة.

تكمن جودة محرك التعرف على الكلام في مدى قدرته على صقل عمليات بحثه في قواعد البيانات لإلغاء النتائج غير الدقيقة وغير المطابقة واختيار النتائج الأكثر مطابقة مع المقاطع الصوتية. وهذا يعتمد على جودة نموذج اللغة والنموذج السماعي، وعلى فعالية الخوارزميات لتحليل الصوت والبحث في قواعد البيانات.

وبتطبيقها العملية المفصلة أعلاه، وفّرت جوجل البحث الصوتي وغيره من التطبيقات والخدمات التي تدعم الإدخال الصوتي باللغة العربية بحيث يستطيع المستخدمون في أي دولة عربية التحدث إلى الهاتف للحصول على نتائج بحث بسرعة وسهولة، وبلهجاتهم المحلية.

واجه فريق تكنولوجيا الكلام في جوجل بعض التحديات الخاصة باللغة العربية وهي تغذية قواعد البيانات بشكل يجعل النظام يفهم تشكيل الكلمات وتمييز لفظها ومعانيها، والفصل بين الكلمات خاصة تلك التي يكون متصلاً بها حروف عطف أو جر أو ضمائر أو غيرها.

لتدريب نظامنا على التعرف على هذه اللغة الفريدة، جمع الفريق أكثر من مليون كلمة ولفظة تستخدمها اللغة العربية المحكية في المناطق ذات الكثافة السكانية الأعلى في كل بلد عربي، وعمل على تدريب النظام بشكل يمكنّه من التعرف على اللهجة الخليجية واللهجات المشرقية واللهجة المصرية وغيرها.

وبالرغم من أنه حين أطلقت الخدمة في البداية، كانت هناك صعوبة في التعرف على كل الكلمات في اللهجات كافة، إلا أنّ ميزة محرك التعرف على الكلام تتمثل في القدرة على تخزين المعلومات في السحاب الإلكتروني، ما يضمن تحسين دقة البحث الصوتي وجودته كلما استخدمه عدد أكبر من الأشخاص.

بالانتقال إلى عملية تطوير الإخراج الصوتي أو ما أطلقنا عليه اسم توليف الكلام أي تحويل النص إلى كلام منطوق، تجدر الإشارة إلى أنّها تشبه العملية المفصلة أعلاه من حيث تحليل اللغة صوتيًّا ولغويًّا والحاجة إلى محرك توليف كلام وقواعد بيانات دقيقة وشاملة، إلى أنها ترتكز بشكل أساسي على سلسلة من جلسات تسجيل لمقاطع وجمل من اللغة يجريها ممثل أو موهبة صوتية في الاستوديو.

تتم عملية تحويل النص إلى كلام من خلال نظام يسمى بمحرك توليف الكلام وهي تتألف من مرحلتين أساسيتين.

تتمثل المرحلة الأولى بتحويل النص الخام الذي يتضمن رموزًا كالأرقام والتواريخ والاختصارات إلى كلمات كاملة مكتوبة ويطلق على هذه العملية اسم "تطبيع النص" (Text Normalization) وهي أشبه بالمعالجة الأولية للنص، يليها عملية تدوين النص باستخدام الأبجدية الصوتية من خلال تحديد النسخة الصوتية التي تطابق كل كلمة. يتم بعد ذلك تقسيم النص إلى وحدات نبرية أي إلى عبارات وجمل وفقرات (حسب الإيقاع والنبرة والتنغيم في الكلام).

أما المرحلة الثانية فتقضي بتحويل التمثيل اللغوي الرمزي للنص إلى صوت وكلام منطوق. لتنفيذ هذه العملية، يرتكز المحرك على قواعد بيانات لمقاطع نصية مسجلة في الاستوديو. وعند إنشاء قاعدة البيانات هذه، يتم تقسيم كل كلام مسجّل إلى مقاطع صرفية (مورفيم) وكلمات وعبارات وجمل.

تضمن قاعدة البيانات هذه الحصول على إخراج صوتي طبيعي عبر اختيار الوحدات في قاعدة بيانات الكلام المسجل ومطابقتها مع وحدات النص المكتوب لإنتاج النسخة المنطوقة من النص الأساسي المكتوب. نظرًا لأنّ عملية توليف الكلام هذه تستخدم قاعدة بيانات لكلام مسجل بصوت شخص حقيقي، فإنها تضمن إخراجًا صوتيًّا طبيعيًّا وعالي الجودة يشبه الكلام البشري بلفظه وإيقاعه ونبرته.

تحديات تكنولوجيا الكلام في اللغة العربية

يتطلب دعم الإدخال الصوتي والإخراج الصوتي في اللغة العربية جهودًا متعددة المجالات والاختصاصات تشمل تحليلاً لفظيًا وصوتيًّا للغة، وتطبيق تقنيات معالجة الكلام العربي، ومعالجة اللغة الطبيعية، وإعداد نظام تعرف على اللغة العربية ونظام توليف اللغة العربية.

تتمثل التحديات الأساسية لإنشاء أنظمة عالية الدقة للتعرّف على الكلام العربي أو توليفه في ما يلي:

- غياب التشكيل في العدد الأكبر من النصوص المكتوبة ما يجعل من الصعب على النظام إعطاء النسخة الصوتية الدقيقة كالتفرقة بين "كَتَبَ" الفعل و"كُتُب" جمع كتاب مثلاً، أو التفرقة بين المؤنث والمذكر، أو المنصوب والمجرور والمرفوع لضمان إخراج صوتي يحترم قواعد اللغة ويعكس المعنى الدقيق للنص المكتوب.
- تعدد وتنوع اللهجات المحلية إلى جانب استخدام اللغة الفصحى ما قد يؤثر على دقة النسخة المكتوبة التي ينتجها نظام التعرف على الكلام، ويزيد من صعوبة التعرف على اللهجة بشكل صحيح.
- تتصل بالكلمات العربية بعض الحروف (حروف الاستفهام مثلاً) أو الضمائر سواء في مطلع الكلمات أم آخرها، الأمر الذي يصعّب على نظام التعرف على الكلام الفصل بين الكلمات بشكل دقيق ونسخها أو نطقها بطريقة صحيحة.
- يزيد استخدام كلمات أجنبية في النصوص العربية المكتوبة أو في الكلام العربي المحكي من صعوبة فصل أنظمة إدخال وإخراج الكلام بين وحدات الكلام العربية والوحدات الأجنبية وبالتالي نسخها أو نطقها بشكل صحيح.

الخاتمة

لا تزال تكنولوجيا الكلام بشكل عام وفي اللغة العربية بشكل خاص مجالاً حديث الولادة ويفتح الأبواب أمام المزيد من الاستكشاف والتطوير. ولكنّ الدعم المتوفر حتى الآن في اللغة العربية في منتجات جوجل، وإن كان لا يزال في بداياته بالمقارنة مع دعم اللغة الإنجليزية أو غيرها، إلا أنه واعد جدًّا. وبالفعل أنّ الإنجازات التي تحققت حتى الآن من القدرة على التعرف على اللهجات كافة في البحث الصوتي مرورًا بالقدرة على الكتابة بالصوت في مستندات جوجل، والاستماع إلى إرشادات التنقل في الخرائط، وصولاً إلى وضع المحاثة في تطبيق الترجمة وما يزيل من حواجز لغوية في التواصل بين الناس، إنما هي إنجازات تعد بدعم أكبر وتضمن تحسينًا وتطويرًا لأداء الأنظمة في المستقبل. لقد أحدث تكنولوجيا الكلام نقلة نوعية في تفاعل المستخدم مع الهاتف أو الكمبيوتر، فهي لم تسهّل الاستخدام وحسب بل أتاحت إلى الأشخاص ذوي الاحتياجات الخاصة أو الأشخاص الأميين الوصول إلى المعرفة والاستفادة من المعلومات، وهذا يصب في صميم مهمة جوجل.