كفاءة نماذج التوليد اللغوي LLMs في اكتشاف الأخطاء النحوية لدى متعلمي اللغة العربية كلغة ثانية وتصويبها: نحو تصميم نظام قائم على تحليل الأخطاء المتكررة.

The Efficiency of Generative Language Models (LLMs) in Detecting Grammatical Errors in Arabic as a Second Language Learners: Towards Designing a System Based on Frequent Error Analysis and Correction

د. عثمان بريحة♥

المُعرّف الرّقميّ للمقال:DOI 10.33705/0114-028-001-002

تاريخ الإرسال: 2025/08/27 تاريخ القبــول: 08 / 11/ 2025 تاريخ النشر: 2026/03/15

ملخّص: تستكشف هذه الدراسة كفاءة النماذج اللغوية (LLMs) في اكتشاف الأخطاء النحوية لدى متعلمي اللغة العربية كلغة ثانية، وذلك من خلال تحليل دقيق لأنماط الأخطاء النحوية المتكررة. كما تهدف إلى تقييم أداء أحدث النماذج اللغوية في تحديد الأخطاء النحوية في نصوص متعلمي العربية كلغة ثانية، وتحليل أكثر الأخطاء تكرارًا وتحديد الأسباب الكامنة وراءها.

ومن أبرز النتائج التي تم التوصل إليها أن بعض النماذج أظهرت تفوقا في رصد الأخطاء مثل (Chatgpt) و(Claude) و(Deepseek) قياسا

♥ جامعة الجزائر.

البريد الإلكتروني: dr.berriha@gmail.com (المؤلف المرسل).

بـ (Gemini) و(Grok) مع ضعف عام في دقة التفسيرات النحوية، كما بينت أن جودة بيانات التدريب تؤثر بشكل كبير على أداء تلك النماذج.

كلمات مفتاحيّة: نماذج لغوية؛ أخطاء نحوية؛ متعلمو اللغة العربية؛ تحليل؛ تصويب أخطاء.

Abstract: This study investigates the efficiency of large language models (LLMs) in detecting grammatical errors among learners of Arabic as a second language, with a particular focus on analyzing recurrent error patterns. It further seeks to evaluate the performance of state-of-the-art models in identifying such errors and to shed light on the underlying causes behind the most frequent ones.

Findings reveal that certain models, notably ChatGPT, Claude, and Deepseek, demonstrated superior performance in error detection when compared to Gemini and Grok. However, the overall accuracy of syntactic explanations remains limited, indicating gaps in deeper grammatical anlysis. The study also highlights the decisive influence of training data quality on model performance, underscoring the need for tailored linguistic resources to enhance reliability in processing Arabic texts.

Keywords: Language Models; Grammatical Errors; Arabic as a Second Language Learners; Error Analysis.

1. مقدّمة: أحدث ظهور نماذج التوليد اللغوي (LLMs) تحولا عميقا في مجال اللسانيات الحاسوبية، وذلك بالنظر إلى قدرتها الفائقة في مجال معالجة اللغات الطبيعية(NLP) خاصة ما تعلق بالترجمة الآلية وتوليد النصوص والكشف عن الأخطاء الواردة فيها. ويبرز بشكل فعال جدا استخدام هذه النماذج

في تحديد وتصحيح الأخطاء النحوية لدى متعلمي اللغة الثانية (L2) كمجال تطبيقي واعد، لكنه لم يحظ بعد بالعناية اللائقة وبشكل كافٍ، خاصة في مجال تعلم اللغة العربية كلغة ثانية، نظرا لما تتميز به هذه اللغة من تراكيب صرفية ونحوية معقدة تطرح تحديات كبيرة لمتعلميها، الأمر الذي يؤدي إلى وقوعهم في أخطاء نحوية متكررة. هذا من جانب، ومن جانب آخر نجد أن التحليل التقليدي للأخطاء يعتمد على التدقيق اليدوي والأنظمة القائمة على القواعد، مما يمكن من دمج النماذج اللغوية الكبيرة وتوفبر منهج أكثر كفاءة وقابلية للتوسع في اكتشاف الأخطاء وتصويبها.

وفي هذا السياق تأتي هذه الدراسة لتقوم بتقييم كفاءة نماذج التوليد اللغوي (LLMs) في اكتشاف الأخطاء النحوية لدى متعلمي اللغة العربية كلغة ثانية بهدف تصميم نظام ذكي لتحليل تلك الأخطاء المتكررة وتقديم تغذية راجعة تجنب المتعلمين الوقوع في تلك الأخطاء مرة أخرى. وتعد كذلك استجابة للحاجة المتزايدة إلى أدوات تعليمية آلية في تعليم اللغات، خاصة في السياقات التي يصعب فيها تقديم التصويبات الفردية. ذلكم ورغم التقدم الكبير في نماذج التوليد اللغوي، فإن أداءها في التعامل مع تعقيدات القواعد العربية – خاصة في النصوص المنتجة من متعلمين – لم يُفحص بشكل كافٍ بعد، وهذا القصور يثير تساؤلات جوهرية حول مدى قدرة هذه النماذج على التكيف مع أنماط الأخطاء الشائعة لدى متعلمي العربية، وإمكانية تفوقها على الطرق التقليدية لتحليل الأخطاء.

وعليه تتمحور إشكالية هذه الدراسة حول تساؤلات رئيسية هي: أولها هل لنماذج التوليد اللغوي القدرة على تحديد وتصنيف الأخطاء النحوية لدى متعلمي العربية بدقة؟، ثانيها هل تتكيف هذه النماذج مع الحساسية التي تفرضها

الخصوصيات الصرفية والنحوية في اللغة العربية؟، وآخرها هل هذه النماذج لها كفاءة في توليد تصويبات مفيدة تربويًا؟.

ولا بد أن نشير هنا إلى أن معظم الدراسات السابقة قد ركزت على اكتشاف الأخطاء باستخدام النماذج اللغوية على لغات كالإنكليزية واللغات الهندوأوروبية الأخرى، بينما ظلت اللغة العربية غير ممثلة بالقدر الكافي. فضلا عن ذلك تعتمد معظم أنظمة التصحيح الآلي الحالية للعربية على قواعد نحوية محددة مسبقًا، والتي قد تفتقد المرونة الكافية لاستيعاب أنماط الأخطاء المعقدة والمتنوعة والمتطورة لدى متعلميها كلغة ثانية.

وللإجابة عن هذه التساؤلات نطرح الفرضيات الآتية:

– يمكن لنماذج التوليد اللغوي تحقيق دقة عالية في اكتشاف الأخطاء النحوية في نصوص متعلمي العربية، متفوقةً على الأنظمة التقليدية القائمة على القواعد؛

– يختلف أداء النماذج حسب نوع الخطأ، حيث تكون أكثر دقة في الأخطاء النحوية مقارنةً بالأخطاء الصرفية نظرًا لتعقيدات الصرف العربي؛

–ضبط النماذج اللغوية على نصوص المتعلمين سيطور قدرتها على اكتشاف الأخطاء وتصحيحها، ويجعلها أكثر تكيفًا مع الأخطاء الشائعة لدى متعلمي العربية.

أما الأهداف التي رصدت في هذه الدراسة فتتمثل في تقييم أداء أحدث النماذج اللغوية في تحديد الأخطاء النحوية في نصوص متعلمي العربية، وتحليل أكثر الأخطاء تكرارًا والأسباب اللغوية الكامنة وراءها، واقتراح إطار عمل لنظام ذكي لتحليل الأخطاء يعتمد على النماذج اللغوية لتقديم تصويبات آنية ومراعية للسياق. منهجيًا، تعتمد الدراسة على نهج متعدد الطرق، يجمع بين التحليل الكمي للأخطاء والتقييم النوعي للتصويبات المولدة آليًا. وسيتم

توظيف مدونة نصوص مكتوبة من قبل متعلمي العربية، مع تدقيقها يدويًا لتحديد الأخطاء النحوية، ومن ثم قياس دقة النماذج اللغوية باستخدام مقاييس الدقة (Precision)، والاستدعاء (Recall)، ودرجة (F1). بالإضافة إلى ذلك سنقوم بتقييم مدى ملاءمة التصويبات المقدمة من النماذج من الناحية التعليمية، ولن تُسلط هذه الدراسة الضوء فقط على نقاط القوة والضعف في أداء النماذج اللغوية في التعامل مع أخطاء متعلمي العربية، بل ستُمهد الطريق أيضًا لتطوير أدوات تعليم لغوي أكثر ذكاءً وتكيفًا.

2.الجانب النظري:

1.2 نماذج التوليد اللغوي (LLMs): تعريفها وتطورها:

أ. تعريف نماذج التوليد اللغوي: أحدث ظهور نماذج التوليد اللغوي (Large Language Models)(LLMs) تطورا ملحوظا في مجال الذكاء الاصطناعي ومعالجة اللغة الطبيعية، حيث أعادت هذه النماذج المتطورة تشكيل حدود النص المُولَّد آلياً، مما أبان عن قدرات فائقة في فهم اللغة البشرية وتوليدها ومعالجتها. وبالاستناد إلى البيانات النصية الضخمة كشفت نماذج مثل سلسلة(GPT) من(Open AI)، و(BERT) من (Google)، و(LLaMA) من (Meta) عن كفاءة عالية في توليد النصوص وترجمتها وفهمها وتلخيصها والانخراط في محادثات تفاعلية مع المستخدمين.

وتعرف نماذج التوليد اللغوي بأنها نماذج ذكاء اصطناعي تقوم على تقنية المحولات(Transformes) يتم تدريبها على كميات هائلة من البيانات النصية كي تصبح قادرة على الفهم الدلالي للغة البشرية ونمذجتها. وقد أظهرت هذه النماذج نتائج مبهرة فيما يتعلق باتساق ودقة سياق النصوص بدرجة تشبه النصوص التي ينشها الإنسان[1] (Ozdemir, 2023).

ب. تطور نماذج التوليد اللغوي: قبل ظهور هذه النماذج القائمة على المحولات سنة 2017م كانت بعض نماذج اللغات تعد ضخمة قياسا بالقيود والمعايير الإحصائية والبيانية في ذلك الوقت، ففي أوائل التسعينيات كانت نماذج (IBM) الإحصائية رائدة في مجال تقنيات محاذاة الكلمات المستعملة في الترجمة الآلية، وهذا ما مهد الطريق لنمذجة اللغة فيما بعد.

والبداية كانت عام 1960م حين ظهر نظام (ELIZA) من قبل جوزيف وايزنباوم، وهو من أوائل روبوتات الدّردشة الآلية التي تقوم باستخدام مطابقة الأنماط، حيث يشتغل أساسا على محاكاة المحادثة الأساسية ويكتشف الكلمات الرئيسية في مدخلات المستخدم ثم يستجيب وفقا لها فيما بعد. ثم وفي سنة 1990م ومع التطور الحاصل وزيادة الاعتماد على شبكة الأنترنيت تم تطوير الشبكات العصبية المتكررة (RNNs) لمعالجة البيانات المتسلسلة كالنصوص والكلام، لكنها واجهت مشكلة مع التسلسلات الطويلة وتم القضاء على هذه المشكلة بإنشاء شبكات الذاكرة طويلة المدى القصيرة (LSTM) فيما بعد.

ومع بداية الألفية الجديدة وزيادة فرص الوصول إلى شبكة الأنترنيت عمل الباحثون على تجميع بينات ضخمة من الويب، وذلك من أجل زيادة حجم بينات التدريب الخاصة بهذه النماذج، ففي سنة 2012م ومع ظهور ما يعرف بالتعلم العميق شهد هذا المجال تحولا عميقا، ولم تحل سنة 2014م حتى عمل الباحثون على إنشاء وحدات متكررة ذات بوابات (GRUs) كنسخة أولية من وحدات (LSTMs) لكنها أبسط وأسرع، الأمر الذي مكن الذكاء الاصطناعي من التركيز على الأجزاء الأهم في تلك التسلسلات الطويلة وفهمها بشكل أفضل.

وفي سنة 2017م تم إنشاء نموذج ذكاء اصطناعي من قبل فريق Google Brain قائم على تقنية التعلم العميق (deep Learning)، واستخدم هذا النموذج

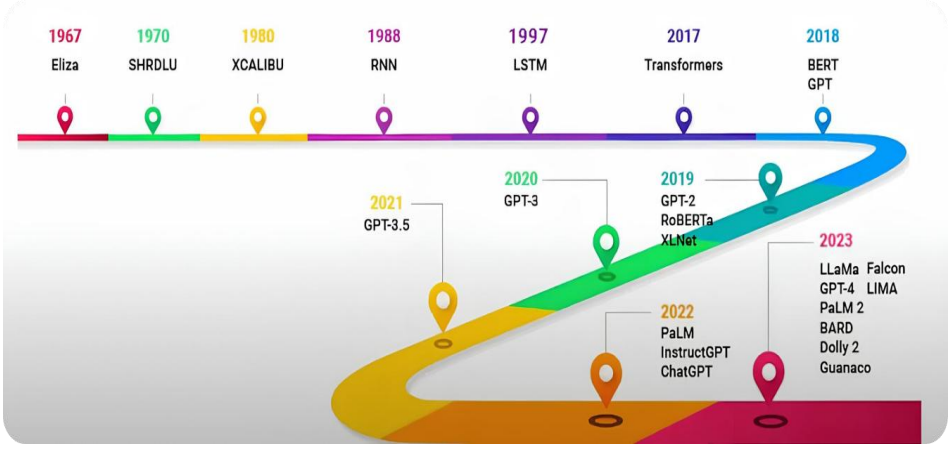
في معالجة اللغة الطبيعية (Natural Language Process) وما يتعلق بها من تحليل النصوص وترجمتها وتحويلها إلى كلام. وفي عام 2018م طورت شركة (Open AI) نموذج المحول التوليدي المدرب مسبقا (Generative Pre-trained Transformer) أو ما يعرف بـ(GPT)، وذلك لإنشاء نصوص تشبه النصوص التي ينشئها الإنسان، وعملت نفس الشركة على تطوير هذا النموذج لسلسلة من الإصدارات المتطورة من حيث عدد معلمات النموذج وحجم البيانات المستخدمة في تدريبه، كما قامت في سنة 2022م بإطلاق (ChatGPT) وهو عبارة عن روبوت محادثة (Chat bot) يستخدم النموذج اللغوي الكبير (GPT) لإنشاء نص يشبه النص الذي ينشئه الإنسان بناء على أوامر مكتوبة[2] (Caelen, O., & Blete, M, 2023).

ويمكن تلخيص مراحل تطور نماذج التوليد اللغوي (**LLMs**) كما ذكرنا في الفقرات السابقة بالخطاطة الآتية التي تبين أهم مفاصل تطور تلك النماذج وأبرز التحديثات التي مرت بها حتى وصلت إلى شكلها النهائي المعروفة به حاليا، وتجدر الإشارة هنا إلى أن هذه الخطاطة تبرز ما وصلت إليه هذه النماذج حتى سنة 2023، إذ لم تتوفر بعد البيانات الخاصة بهذه النماذج في سنتي 2024م و2025م بعد:

الشكل 1: تطور نماذج التوليد اللغوي (LLMs)

المصدر[3]: (فينا، 2024)

وهذه النماذج تم تدريبها مسبقا على كميات هائلة من البيانات النصية فهي في الأساس نماذج شبكة عصبية تعالج وتولد نصوصا، وقد تم تدريبها على الأقل على مليار كلمة حيث تقوم باستنتاجات بناء على التعلم الانتقالي (Transfer Learning)، ويبين الجدول أدناه حجم البيانات التي توجد في كل



نموذج من النماذج الثلاثة: (ELMO)، (BERT) و(GPT) وبيانات التدريب الخاصة بها:

الجدول 1: طبيعة الشبكة العصبية لنماذج التوليد اللغوي LLMs وعدد المعلمات وحجم بيانات التدريب

Model	Architecture	Objective	Params #	Training Data	Papers
ELMO	BiLSTM	Autoreg.LM	94M	1BWords	I
BERTbase	Tf.Encoder	MLM/NSP	110M	3.3Words	IV-I
BERTlarge	Tf.Encoder	MLM/NSP	340M	BWords3.3	-
GPT-2	Tf.Decoder	Autoreg.LM	1.5B	9BTokens	IV
GPT-3	Tf.Decoder	Autoreg.LM	175B	300BTokens	-
GPT-4	?	?	?	?	V

المصدر[4]: (Kunz, 2024)

هذا الجدول يبين مسار التطور السريع الذي شهدته نماذج التوليد اللغوي، سواء من حيث البنية أم الحجم أم الهدف التدريبي، كما يبرز التحديات المتزايدة المتعلقة بمتطلبات البيانات والقدرة الحوسبية. ويُعد هذا العرض المختصر قاعدة جيدة لفهم الفروقات التقنية بين النماذج وللمقارنة بين توجهات التصميم المختلفة التي اعتمدها الباحثون في سبيل تحسين الأداء اللغوي لهذه النماذج الحديثة. وعليه فهو يحتوي على مقارنة موجزة بين عدد من نماذج التوليد اللغوي الكبرى (LLMs) من حيث البيانات التدريبية، وعدد المعلمات (Parameters)، والهدف التدريبي، والبنية المعمارية، والنموذج المستخدم. وقد تم تصنيف النماذج تصاعديًا بحسب حجمها وتطورها الزمني بدءًا من "ELMO"وصولًا إلى."GPT-4" .

ويُلاحظ أن النماذج الأولى مثل(ELMO) وبنية(BERT) بنسختيها (base) و(large) اعتمدت على أهداف تدريبية مثل نمذجة اللغة المقنعة (MLM) وتوقع الجملة التالية NSP، مع استخدام معمارية المشفِّر (Transformer) أما النماذج الأحدث مثل GPT-2 وGPT-3 فقد اعتمدت على نمذجة اللغة التوليدية (Autoregressive LM) باستخدام معمارية المفكك

(Transformer)، مع ارتفاع كبير في عدد المعلمات وحجم البيانات التدريبية. يعكس هذا الانتقال تغيرًا نوعيًا في تصميم النماذج، إذ اتجه الباحثون نحو بنى أكثر قابلية للتوسّع وذات قدرات تعميم أعلى.

2. تعليمية اللغة العربية كلغة ثانية: تنامى –في العقود الأخيرة– الاهتمام بتعلم اللغة العربية بشكل ملحوظ جدا وذلك لعدة عوامل تشمل مكانة هذه اللغة بوصفها لغة للقرآن الكريم والسنة النبوية وما يتصل بهما من شرائع وأحكام وكونها نافذة لفهم التراث العلمي والثقافي العربي والإسلامي، ناهيك عن أنها إحدى اللغات الست الرسمية في الأمم المتحدة، فضلا عن التوسع الاقتصادي الذي عرفته الدول العربية مؤخرا. فتعلم اللغة العربية ليس عملية نقل معرفي فقط بل هو تسهيل للتواصل الفعال في عالم يزداد ترابطا يوما بعد يوم.

واللغة العربية بخصائصها اللغوية الفريدة تمثل تحديا لمتعلميها من لغات أخرى خاصة ما تعلق بالتعقيدات في بنيتها الصرفية والنحوية كنظام الإعراب والاشتقاق والتأنيث والتذكير والجموع غير القياسية وغيرها، الأمر الذي يتطلب جهدا مضاعفا لاتقانها. لذلك يتطلب تعليم اللغة العربية كلغة ثانية استراتيجيات تدريسية خاصة تقوم على تبسيط قواعدها الصرفية والنحوية، والاستعانة بالتكنولوجيات الحديثة كالوسائط المتعددة والتطبيقات التفاعلية لتجاوز هذه التحديات، وجعل عملية تعلم هذه اللغة عملية أكثر سلاسة وفعالية.

2.2 نظريات تعليم اللغة لغير الناطقين بها: عُرفت في هذا المجال العديد من النظريات التي دعت إلى تطوير أساليب وطرائق التدريس المتبعة في تعليم اللغة لغير الناطقين بها، وأهم هذه النظريات هي:

1. نظرية التطابق: هذه النظرية تقوم على مبدأ أن اكتساب اللغة الأم وتعلم اللغة الثانية عمليتان متطابقتان، وينفي أصحابها تأثير اللغة الأم على تعليم اللغة الثانية، وعند تطبيق هذه النظرية يلاحظ وجود فروق كبيرة من النواحي

اللغوية والنفسية بين المتعلم الطفل والمتعلم البالغ، حيث يتقن المتعلم البالغ لغة واحدة على الأقل هي لغته الأم، وقد أثبتت دراسات تأثيرا جزئيا للغة الأم على اللغة الثانية من خلال الأخطاء التي يرتكبها المتعلم[5] (خرما، نايف وحجاج علي، 1988م).

2. نظرية التحليل التقابلي (التقابل اللغوي): هذه النظرية ظهرت كرد على النظرية السابقة نظرية التطابق، ورائدها روبرت لادو اللغوي الأمريكي الذي نادى بإجراء مقارنة بين اللغة الأم واللغة الثانية، حيث تشمل هذه المقارنة التشابه والاختلاف في الأصوات والقواعد اللغوية، وذلك للتنبؤ بالصعوبات التي يمكن أن تواجه متعلمي اللغة الثانية[6] (الراجحي، 2000م).

3. نظرية تحليل الأخطاء: تقوم هذه النظرية على مبدأ تحليل الأخطاء اللغوية التي يقع فيها متعلمو اللغة الثانية، حيث يتم تحديدها وتحليلها وتفسيرها والتنبؤ بجوانب القصور ومحاولة معالجتها لدى المتعلمين، وهي أكثر واقعية وموضوعية من نظرية التحليل التقابلي؛ إذ تعالج أخطاء حقيقية بينما تعالج الأخرى أخطاء متوقعة أو ما يتنبأ به من أخطاء قد لا يقع فيها المتعلمون، وتقوم هذه النظرية على ثلاث مراحل لتحليل الأخطاء هي: تحديد الأخطاء ووصفها، ثم تفسير تلك الأخطاء، وأخيرا تصويبها وعلاجها، ويعتمد في ذلك على النموذج النحوي لأنه النموذج النظري القادر على وصف الخطأ، ليتبين لمتعلم اللغة كيف خالف قواعد التحقيق عند صياغة الجملة[7] (صيني 1982م).

3. الدراسة التطبيقية: تحدثنا سابقا في الجزء النظري عن أهم المفاهيم النظرية المتعلقة بهذه الدراسة ممثلة في التعريف بنماذج التوليد اللغوي وتطورها، ثم عرض لأهم النظريات الخاصة بتلك اللغة لغير الناطقين بها مثل نظرية التطابق ونظرية التحليل التقابلي ونظرية تحليل الأخطاء، هذه الأخيرة

التي ستكون مرتكز هذه الدراسة التطبيقية إذ تتناول تحليل الأخطاء النحوية؛ وذلك لأن النموذج النحوي هو النموذج النظري القادر على وصف الأخطاء.

وفي هذا الجزء من الدراسة سنقدم الخطوات المنهجية والإجرائية التي سنتوصل من خلالها إلى قياس كفاءة نماذج التوليد اللغوي في تحديد الأخطاء النحوية وتصويبها لدى متعلمي اللغة العربية كلغة ثانية، وتشمل هذه الخطوات ما يلي:

1- عرض التطبيقات التي ستجرى عليها التجربة.

2- تحديد عينة الدراسة.

3- إجراء التجربة: صياغة التعليمة وتحديد معايير التقييم.

4- عرض نتائج التجربة وتحليلها.

عرض التطبيقات التي ستجرى عليها التجربة:

1. 1. (ChatGPT) يعتمد (ChatGPT) الذي قدمته شركة (Open AI) أول مرة عام 2018م على نموذج لغة المحول المدرب مسبقا GPT، وهو نموذج ذكاء اصطناعي يعتمد على التوليد اللغوي وتقنيات التعلم العميق. ويتمتع (ChatGPT) بقدرة فائقة على إنتاج لغة تشبه اللغة البشرية بدءا من الإجابة على الاستفسارات البسيطة وصولا إلى المهام الأكثر تعقيدا مثل إنشاء خطابات التقدير وتسهيل المحادثات الصعبة مع الكم الهائل من بيانات النصوص والمدخلات المتنوعة والمتاحة[8].(Gupta & Hathwar, 2020)

وفي هذه الدراسة سنعتمد على آخر تحديث لـCatGPT وهوChaGPT-5 الذي تم طرحه مؤخرا بتاريخ 2025/08/08م بعد خضوعه لمجموعة من التحسينات ليكون أكثر دقة وقابلية للتكيف، وأكثر توافقا مع النوايا البشرية.

1. 2. Claude: هو أحد نماذج التوليد اللغوي واسعة النطاق تم تطويره من قبل شركة Anthropic صدر أول مرة في شهر مارس 2023م، وهو نموذج

متعدد الوسائط يقوم بمهام عديدة كالإجابة عن الأسئلة والبحث والتدقيق اللغوي والتحرير وتلخيص المستندات (PDF,word) وإنشاء النصوص والمحتوى والترجمة وإنشاء خطط العمل ومعالجة الصوت والصورة، وأهم ما يميز (Claude) هو الكفاءة العالية حيث يستطيع معالجة حوالي 30 صفحة نصية في الثانية الواحدة وقراءة أوراق وبحثية في أقل من ثلاث ثوان أي أسرع مرات من نظرائه[9]. (Grammarly,2024).

1. 3. **Gemini:** يعد (Gemini) النموذج اللغوي الكبير (LLM) الخاص بشركة (Google) وكان يعرف سابقا باسم (Project Bard) وهو مصمم خصيصا لمعالجة طرائق وأنواع متعددة من البيانات بما في ذلك الصوت والصور ورموز البرامج والنصوص والفيديو، كما يشغل (Gemini) روبوت المحادثة المدعوم بالذكاء الاصطناعي التوليدي، وقد تم تدريبه على مجموعة ضخمة من البيانات متعددة اللغات والوسائط حيث يستخدم نموذج المحول وهو بنية شبكة عصبية قدمتها (Google)[10]. (IBM,2025).

1. 4. **Grok:** تم إطلاق (Grokc3) في 17 فبراير 2025م، وهو نموذج لغوي كبير (LLM) تم تطويره بواسطة (Xai). و(Grok) هو مساعد مدعوم بالذكاء الاصطناعي يساعدك في إكمال المهام، مثل الإجابة عن الأسئلة وحل المشكلات والعصف الذهني. يتوفر (Grok) لمستخدمي X ويتم تشغيله بواسطة نموذج اللغة الكبير (LLM) المتطور من (Xai)، حيث يقدم (Grokc3)، الذي يستند إلى هذا الأساس، طريقة تفكير محسّنة وأوقات استجابة أسرع ووصولاً فورياً إلى المعلومات. على غرار إصداراته السابقة، تم دمج (Grokc3) مع X (تويتر سابقًا)[11]. (help.x, 2025)

1. 5. **Deepseek:** هو نموذج قدمته شركة (Deepseek) الصينية في 2025/01/20م ويحمل اسم (Deepseek-R1-0528) وهو نسخة محدثة

ومفتوحة المصدر وبتكلفة زهيدة قياسا بالنماذج الأخرى، وقد عملت الشركة على تطوير نماذجها وتحديثها منذ تأسيسها عام 2023م حيث طرحن نموذج (Deepseek Coder) في نوفمبر 2023م، ثم أصدرت نموذج (Deepseek LMM) في ديسمبر 2023م وهو نموذج أولي متعدد الأغراض، وفي ماي 2024م تم إصدار (Deepseek-2) وهو الإصدار الثاني من برنامج (LMM) وفي نفس العام أي في جويلية 2024م طرحت الشركة نموذج (Deepseek-Coder-V2) الخاص بالبرمجة المعقدة، وآخر نموذج تم إصداره في ماي 2025 هو (Deepseek-R1-0528) وهو نسخة محدثة من (R1) الأصلي ويتميز بانخفاض معدلات الهلوسة وعمق كبير في التفكير المنطقي[12].(techtarget, 2025).

2. تحديد عينة الدراسة: تكتسي عملية تحديد عينة الدراسة أهمية بالغة وتعد خطوة أساسية لضمان دقة النتائج وقابليتها للتعميم، وينبغي أن تكون ممثلة للفئة المستهدفة من حيث المستوى اللغوي والخلفية التعليمية وتنوع الأخطاء؛ لأن ضبط شروط العينة وخصائصها سيسهم في توفير بيانات موثوقة تسمح بتقييم أداء نماذج التوليد اللغوي في اكتشاف الأخطاء وتصحيحها على نحو موضوعي، وسيساعد على الكشف عن جوانب القوة والقصور فيها ويحدد مدى التفاوت الحاصل بينها في معالجة الأخطاء وتحديد أسبابها، هذا من جهة، ومن جهة أخرى سيمنح نتائج الدراسة مصداقية علمية، ويزيد من فرص الإفادة منها في عملية تطوير استراتيجيات تعليمية وتقويمية مبتكرة.

بناء على ما سبق فإن عينة البحث يجب أن تتوفر فيها الشروط الآتية:

– تنوع مصادر الأخطاء النحوية وملاءمتها لفئة متعلمي اللغة اللعربية كلغة ثانية، واستجابة لهذا الشرط تم اختيار خمسين (50) جملة من الأبحاث المتخصصة في تحليل الأخطاء لدى هذه الفئة من المتعلمين؛

– يجب أن يكون حجم العينة (50 جملة) كافيا لتمثيل مختلف الأخطاء النحوية الشائعة لدى فئة المتعلمين وملائما لأهداف الدراسة ومتطلباتها؛

– تنوع الأخطاء النحوية وذلك بغرض قياس كفاءة النماذج اللغوية على تحديد الأخطاء وتصحيحها.

والجدول المرفق يضم الجمل التي تم اختيارها بناء على الشروط المحددة سابقا:

الجملة	الخطأ	الجملة	الخطأ
الرجل يركب في ظهر الفرس	في ظهر الفرس	مكة والمدينة مكان آمن	مكان آمن
هي اللغة القرآن الكريم	اللغة القرآن الكريم	حياة في الجزائر جيدة	حياة في الجزائر
وصلت إلى جامعة الإسلامية	جامعة الإسلامية	تمرنت على القراءة صامتة وجهرية	القراءة صامتة وجهرية
كان جو باردا	جو باردا	أكتب لك هذه رسالة	هذه رسالة
في المكة موجودة عرفة ومنى	في المكة	أصعب شيء في اللغة العربية هو تحدث	هو تحدث
الدموع يسيل على خديها	يسيل على خديها	قبل أنتقل	أنتقل
أكلت الولد سمكة	أكلت الولد	كلهم يريدوا أن يتكلمون الإنكليزية	يريدوا أن يتكلمون
وجدنا الطقس فيها جميلة	الطقس فيها جميلة	أتوقع أن أريك	أن أريك
أنا ذهبت كل يوم إلى الجامعة	أنا ذهبت كل يوم	أنت تدرسي جيدا	تدرسي جيدا
ذهبت إلى المطعم قريب من الفندق	قريب من الفندق	لا بد تشاهدي	تشاهدي
زرت خالتي وبتُّ في شقتهم	وبتُّ في شقتهم	أريد أعمل في الجزائر	أريد أعمل
سأزور إلى كل دول العربية	إلى كل دول العربية	ولكن أرغب أن	ولكن أرغب أن
تختلف القواعد العربية من الإنكليزية	القواعد العربية من الإنجليزية	سأستمر دراسة اللغة العربية	دراسة اللغة العربية

أقرأ كتاب عربي	كتاب عربي	لا أفهم كثير من المفردات	كثير من المفردات
ليس جيد	ليس جيد	لذلك لا أثق بنفس	أثق بنفس
هناك مشاكل الدراسة اللغة العربية	الدراسة اللغة العربية	العواصف الكثيرة في مدينتك	الكثيرة في مدينتك
يصور مختلف حيوانات	مختلف حيوانات	الترويح عن النفس مفيدة	عن النفس مفيدة
القراءة مفيد جدا للعقل	مفيد جدا للعقل	لم يكن في مكان ماء	مكان ماء
يريدون السفر في بلدان مختلفة	في بلدان مختلفة	في قراءة الكتب القصص	الكتب القصص
يحضرون معهم أنواع الأطعم	أنواع الأطعم	ذهبت أقاربي لأزوره	لأزوره
رجعت في الغد إلى البيت	في الغد إلى البيت	أنا كان في المنزل	كان في المنزل
في حقل التعليم اللغة العربية	التعليم اللغة العربية	له دور كبير وحيويا في العملية التربوية	له دور كبير وحيويا
ما أثبته الأخصائيين	أثبته الأخصائيين	جعلنا لهذا المحور ثلاث سؤالات	ثلاث سؤالات
لمن ذلك السيارة؟	ذلك السيارة	مررت بمطعماً جميلاً	بمطعماً جميلاً
يحب المؤمنين العمل الصالح	يحب المؤمنين	أحب السير في الشوارع الواسع	في الشوارع الواسع

3. إجراء التجربة (صياغة التعليمة وتحديد معايير التقييم): بعد تحديد عينة الدراسة وحصرها يأتي الدور على إجراء التجربة لقياس كفاءة النماذج اللغوية وقدرتها على اكتشاف الأخطاء اللغوية وتصويبها، وسيتم ذلك من خلال صياغة التعليمة المناسبة الموجهة للنماذج اللغوية وتحديد معايير التقييم؛ لأن هذه النماذج مدربة على إنتاج مخرجات معينة بناء على مدخلات محددة، أي انها تشتغل وفق طريقة التعلم القائم على المحفزات، ونقصد بالمحفز نص

التعليمة الدقيق الذي يقدم للنموذج اللغوي ويوجه استجابته بطريقة معينة ويشترط في نص التعليمة أن يكون:

- دقيقا وواضحا بالقدر الكافي؛
- تحديد المهمة المطلوبة من النموذج اللغوي بدقة؛
- توفير السياق الملائم الذي يساعد على فهم التعليمة (المحفز)؛
- تحديد نمط الإجابة وتوضيح شكلها.

ووفقا للشروط الأربعة التي تم تحديدها في التعليمة الموجهة للنماذج اللغوية تم صياغة التعليمة (الأمر المحفز) على النحو الآتي: عالج الجملة (...) معالجة لغوية دقيقة وفق القواعد النحوية للغة العربية الفصحى، وحدد الأخطاء الواردة فيها ثم فسّر سبب الخطأ واذكر التصحيح المناسب، بحيث يكون عرضك بالنمط الآتي: الجملة الأصل، موضع الخطأ، العلة، الصواب. وتجدر الإشارة إلى أن هذه التعليمة في شكلها النهائي قد راعت المؤشرات الكمية والنوعية ومعايير التقييم المحددة، ونقصد بالمؤشرات الكمية نسبة الدقة والاسترجاع في كل نموذج على حدة، أما المؤشرات النوعية فتمثل ملاءمة التصويب وسياق استخدامه. أما معايير التقييم فقد حددت بناء على محددات الدراسة فيما يلي: القدرة على اكتشاف الأخطاء، دقة التصحيحات المقدمة ودقة تفسير الخطأ.

وبعد استكمال التجربة على النماذج اللغوية الخمسة احتكاما للمعايير التي ذكرت سابقا، أفضت هذه التجربة عن نسب متفاوتة في أدائها يلخصها الجدول الآتي:

معايير التقييم / النموذج اللغوي	دقة اكتشاف الأخطاء	دقة التصحيحات	دقة التفسير
CatGPT-5	%94	%85	%65
Claude	%93	%90	%85

Gemini	%60	%75	%50
Grok	%90	%90	%55
Deepseek	%92	%84	%70

4. عرض نتائج التجربة وتحليلها: من خلال التجربة التي أجريت على نماذج التوليد اللغوي تم تسجيل الملاحظات الآتية:

– سجلنا تقدما واضحا لنماذج (Chatgpt) و(Claude) و(Deepseek) على النموذجين المتبقيين (Gemini) و(Grok) في دقة اكتشاف الأخطاء النحوية، حيث بلغت لدى (Chatgpt) نسبة 94% ثم يليه (Claude) بنسبة 93% و (Deepseek) بنسبة 92%، أما (Grok) فبلغت دقة تصحيحه للأخطاء 90%، وحلّ (Gemini) أخيرا بنسبة أقل بكثير من نظرائه بلغت 60%؛

– في دقة التصحيحات المقدمة يتفوق كل من (Claude) و(Grok) على (Chatgpt) و (Deepseek) حيث سجل نسبة تقدر بـ 90% لكليهما، بينما سجل (Chatgpt) و(Deepseek) نسبتي 85% و84%، بينما يبقى (Gemini) متأخرا على نظرائه بنسبة بلغت 75%؛

– في تفسير الأخطاء النحوية يتفوق Claude بشكل واضح عن نظرائه بنسبة بلغت 85%، ثم يليه (Deepseek) بنسبة 70% و(Chatgpt) بنسبة 65%، ويحل أخيرا كل من (Grok) بنسبة 55% و(Gemini) بنسبة 50%؛

– يلاحظ أن (Gemini) له قدرة على تقديم تصحيحات ملائمة أكبر من قدرته على اكتشاف الأخطاء النحوية، حيث بلغت قدرته على التصحيح نسبة 75%، بينما قدرته على اكتشاف الأخطاء لم تتجاوز 60%، وهذا يفسر بأن أخطاء هذه النماذج هي أخطاء أداء وليست أخطاء قدرة، أي أن الفشل في

اكتشاف موضع الخطأ لكنه قدّم تصحيحا ملائما، لذلك هذه النماذج لها قدرة على تحرير النصوص لكنها لا تمتلك القدرة على اكتشاف الأخطاء وتحليلها؛

– على الرغم أن كبر حجم بيانات التدريب الخاصة بهذه النماذج اللغوية إلا أن أداءها لم يكن دقيقا 100% في اكتشاف الأخطاء النحوية وتصحيحها وتفسيرها، ولعل مرد ذلك الأخطاء التي تقع أثناء عملية إعداد البيانات وترميزها من قبل المختصين، أو تكون الأخطاء راجعة إلى النموذج اللغوي نفسه بسبب الصعوبة التي يواجهها في تعلم بعض الأنماط اللغوية التي تميز اللغة العربية عن باقي اللغات، فيصبح عاجزا عن إدراكها وفهمها تقنيا مثل فهم السياق وعلاقة الكلمات ببعضها البعض، فيؤدي ذلك إلى أخطاء تقنية لا يمكن تجنبها؛

– تفاوت أداء النماذج اللغوية في عمليات اكتشاف الأخطاء النحوية وتصحيحها وتفسيرها يعود إلى أمرين: الأول هو الطبيعة الاحتمالية للنماذج حيث تتضمن هذه النماذج علاقات احتمالية تعتمد على الكلمات الأكثر شيوعا في السياقات المختلفة، والثاني هو ضعف التفكير المنطقي للتطبيقات حيث تقدم إجابات عالية الدقة حينما يتطلب الأمر استرجاع معلومات مباشرة لكنها تواجه صعوبة في التعامل مع المسائل التي تتطلب تحليلا أو تفسيرا ما لذلك لاحظنا أن نسبة تفسيرها للأخطاء أقل من قدرتها على اكتشافها وتصحيحها.

4. خاتمة: تناولت هذه الدراسة كفاءة نماذج التوليد اللغوي في تشخيص الأخطاء النحوية لدى متعلمي اللغة العربية كلغة ثانية، مبيّنة قدرتها على معالج تلك الأخطاء بدقة ومرونة عالية. كما أبرزت نتائج الدراسة أن هذا المسار من البحث يفتح آفاقا واعدة في مجال تطوير أنظمة تعليمية ذكية تقوم على دمج النماذج اللغوية في بيئات تعليمية موجهة، ومن ثمّ تحسين جودة

التعليم وتقليص الأخطاء اللغوية بصورة واضحة، ومن أبرز ما توصلت إليه الدراسة نذكر ما يلي:

– قامت هذه الدراسة بتقييم كفاءة نماذج اللغة التوليدية الكبيرة (LLMs) في اكتشاف الأخطاء النحوية لدى متعلمي اللغة العربية كلغة ثانية وتصويبها وذلك من خلال تحليل أداء خمسة من أبرز النماذج اللغوية، وهي(Chatgpt, Claude, Gemini, Grok, وDeepseek)، بهدف الوصول إلى تصميم نظام ذكي قادر على تحليل الأخطاء المتكررة وتقديم تغذية راجعة فعّالة للمتعلمين؛

– اعتمدت الدراسة على منهج تجريبي يقوم على تحليل عينة مكونة من خمسين (50) جملة تضم أخطاء نحوية جُمعت من أبحاث متخصصة في تحليل أخطاء متعلمي العربية كلغة ثانية، ثم توجيه تعليمة لكل نموذج من النماذج الخمسة لمعالجة الجمل، مع تحديد معايير دقيقة للتقييم شملت: دقة النموذج على اكتشاف الأخطاء وقدرته على تحديد الخطأ في الجملة، ودقة التصويبات المقترحة ومدى صحتها، ودقة النموذج في تفسير الخطأ النحوي وشرح سببه بشكل صحيح؛

– أظهرت نتائج الدراسة التطبيقية تفاوتا واضحا في أداء النماذج اللغوية في معالجة الأخطاء النحوية، إذ تتفوق نماذج (Chatgpt) و(Claude) (Deepseek) في اكتشاف الأخطاء النحوية مقارنة بنظرائها (Gemini) وGrok؛

– تتصف النماذج اللغوية بأدائها غير المتسق فرغم قدرتها على تصحيح الأخطاء النحوية إلا أنها تواجه صعوبة في تقديم تفسيرات دقيقة لأسبابها، مما يشير إلى أن قدرتها على التحليل اللغوي العميق لا تزال ضعيفة نسبيا ومحدودة؛

– تتأثر النماذج اللغوية بدقة وجودة بيانات التدريب الخاصة بها حيث تؤثر بشكل واضح جدا على أدائها، مما يدل على أن الأخطاء في بيانات التدريب يؤدي إلى قصور في أدائها؛

– كشفت الدراسة عن الطبيعة الاحتمالية للنماذج اللغوية وأن قدرتها على فهم السياق والتفكير المنطقي لا تزال في مرحلة البناء والتطوير؛

–تشكل اللغة العربية الفصحى تحديا كبيرا بالنسبة للنماذج اللغوية وذلك بالنظر إلى تعقيداتها النحوية وأهمية السياق في إنتاج وفهم اللمتتاليات النصية.

واستنادا إلى نتائج هذه الدراسة يمكن تقديم المقترحات والتوصيات الآتية:

– ضرورة الاستفادة من النماذج اللغوية في تطوير أدوات التصحيح النحوي مع السعي إلى تحسين قدرتها مع النعقيدات اللغوية في اللغة العربية الفصحى؛

– استخدام النماذج اللغوية كأداة مساعدة لمتعلمي اللغة العربية لتقديم تغذية راجعة سريعة للطلاب، مع الأخذ بمحاذير استخدامها ومراجعة ما تطرحه من تصويبات وتفسيرات؛

– تطوير النماذج اللغوية بشكل يراعي متطلبات اللغة العربية وخصوصيتها وتغذيتها بمدونات نصية تحتوي على أخطاء شائعة لدى متعلمي اللغة العربية كلغة ثانية، وذلك بغرض تحسين قدرتها على فهم الأخطاء النحوية وتصويبها؛

– توسيع نطاق البحث ليشمل عينات أكبر وأكثر تنوعا للأخطاء اللغوية كالصوتية والصرفية والدلالية والبلاغية من نصوص المتعلمين، وإجراء دراسات معمقة تكشف عن أداء النماذج اللغوية من جهة وتطوير قدرتها على معالجة تلك الأخطاء من جهة أخرى.

5. قائمة المراجع:

- **المؤلفات:**

- Ozdemir, S, Quick start guide to Large Language Models ; Strategies and best Practices for using Chat GPT and others LLMs, Addison-Wesley Professional, (USA: O'reilly Media, 2023), p 23.
- Caelen, O & Blete, M, Developing Apps with GPT-4 and Chat GPT Addison-Wesley Professional, (USA: O'reilly Media, 2023), p35.
- Kunz, Jenny, Understanding Large Language Models: Towards Rigorous and Targeted Interpretability Using Probing Classifiers and Self-Rationalisation, Linköping University, (Sweden: Linköping University, 2024), p11.
- خرما، نايف وحجاج، علي، اللغات الأجنبية: تعليمها وتعلمها (الكويت: سلسلة عالم المعرفة، 1988م)، ص 33.

- الراجحي، عبده، علم اللغة التطبيقي وتعليم العربية، (مصر: دار المعرفة الجامعية، 2000م)، ص46.
- صيني، محمود، التقابل اللغوي وتحليل الأخطاء، (المملكة العربية السعودية: جامعة الملك سعود، 1982م)، ص 05-06.
- المقالات:

- Aishwar, Gupta & Divya, Hathwar, Introduction to AI Chatbots, International Journal of Engineering Research & Technology (IJERT), Vol. 9 Issue 07, July-2020, p255.

- مواقع الانترنيت:
- فينا أبيرامي (2024م+)، من الرمز إلى المحادثة: كيف يعمل برنامج https://ww.ultralytics.com/ar/blog/from-code-to- LLM conversation-how-does-an-llm-work (Consulté le 13/07/2025).
- Grammarly,(2024) , Claude AI 101: What It Is and How It Works, (Consulté le 11/08/2025).
- IBM, (2024), What is Google Gemini, (Consulté le 11/08/2025).
- help.x, (2025), About Grok, the AI-powered assistant on X, 0 (Consulté le 11/08/2025).
- Techtarget, (2025), DeepSeek explained: Everything you need to know. https ://www.techatarget.com/whatis/feature/DeepSeek-explained-Everything-you-need-to-know (Consulté le 11/08/2025).

6. هوامش:

[1]- Ozdemir, S, Quick start guide to Large Language Models; Strategies and best Practices for using Chat GPT and others LLMs, Addison-Wesley Professional, (USA: O'reilly Media, 2023), p 23.

[2]- Caelen, O & Blete, M, Developing Apps with GPT-4 and Chat GPT Addison-Wesley Professional, (USA: O'reilly Media, 2023), p35.

[3]- https://www.ultralytics.com/ar/blog/from-code-to-conversation-how-does-an-llm-work (Consulté le 13/07/2025)

[4] - Kunz, Jenny, Understanding Large Language Models: Towards Rigorous and Targeted Interpretability Using Probing Classifiers and Self-Rationalisation, Linköping University, (Sweden: Linköping University, 2024), p11.

[5]– خرما، نايف وحجاج، علي، اللغات الأجنبية: تعليمها وتعلمها، (الكويت: سلسلة عالم المعرفة، 1988م)، ص 33.

[6]– الراجحي، عبده، علم اللغة التطبيقي وتعليم العربية، (مصر: دار المعرفة الجامعية 2000م)، ص46.

[7]– صيني، محمود، التقابل اللغوي وتحليل الأخطاء، (المملكة العربية السعودية: جامعة الملك سعود، 1982م)، ص 05–06.

[8] - Aishwar, Gupta & Divya, Hathwar, Introduction to AI Chatbots, International Journal of Engineering Research & Technology (IJERT), Vol. 9 Issue 07, July-2020, p255.

[9]- https://www.grammarly.com/blog/ai/what-is-claude-ai / (Consulté le 11/08/2025).

[10]- https://www.ibm.com/sa-ar/think/topics /google-gemini (Consulté le 11/08/2025)

[11]- https://www.ibm.com/ ar/using-x/about-grok (Consulté le 11/08/2025).

[12] - https://www.techtarget.com/ whatis/feature/deepSeek-explained-Everything-you-need-to-know (Consulté le 11/08/2025).