Interdisciplinary Studies in Arabic Language and Literature

Online ISSN: 3092-6955

A Morphological and Syntactic Study of Neologisms and Structures in Arabic Social Media: Twitter, Facebook, Instagram as Case Studies

Soudabeh mozaffari[1],Ramin sedaghat*[2]

[1] Associate Professor in Arabic Language and Literature, Faculty of Literature and Humanities, Kharazmi University, Tehran, Iran.

[2] MA student of Arabic language and literature, Faculty of Literature and Humanities, Kharazmi University, Tehran, Iran.

Article Info

Article type:
Research Article

Received:
28/02/2026

Accepted:
11/05/2026

ABSTRACT

The digital revolution has created a new linguistic arena on social media, characterized by a hybrid mix of Modern Standard Arabic, colloquial dialects, and foreign languages. Adopting a mixed descriptive-analytical and quantitative approach, this study analyzes a purposive sample of 500 public posts from X, Facebook, and Instagram (2025). Morphologically, the findings reveal the prevalence of "Arabizi" (42%), alongside hybrid derivation from foreign roots and blending. Syntactically, a strong simplification trend is observed, marked by the substitution of standard relatives with the unified colloquial *illi*, and a stylistic shift towards SVO structure (65% of declarative sentences). While these transformations serve as identity negotiation tools, they pose significant structural challenges, causing up to 30% semantic failure in generative AI models. This interdisciplinary study highlights the urgent need for annotated hybrid corpora to enhance Arabic NLP systems.

Keywords: *Morphological and Syntactic Transformations, Neologisms, Social Media Language, Arabizi, Digital Linguistics, Artificial Intelligence, NLP.*

Cite this article:mozaffari,S. & sedaghat,R. (2026). A Morphological and Syntactic Study of Neologisms and Structures in Arabic Social Media: Twitter, Facebook, Instagram as Case Studies, *Interdisciplinary Narrative Studies in Arabic Language and Literature*. 3(Special Issue). 93-118.

DOI: 10.22034/jisall.2026.555190.1101

© The Author(s).

Publisher: University of Zabol

***Corresponding Author**: ramin sedaghat

Address: master student of Arabic language and literature, Faculty of Literature and Humanities, Kharazmi University, Tehran, Iran.

E-mail: ramin.sedaghat@khu.ac.ir.

مجلة نصف سنوية دراسات متعددة التخصصات في اللغة العربية وآدابها

الترقيم الدولي الالكترونيكي: ٦٩٥٥-٣٠٩٢

دراسة صرفية نحوية للألفاظ والتراكيب الجديدة في وسائل التواصل الاجتماعي العربية: تويتر، فيسبوك، إنستغرام أنموذجا

سودابه مظفري[١]، رامين صداقت[٢*]

[١] أستاذة مشاركة، قسم اللغة العربية وآدابها، كلية الآداب والعلوم الإنسانية، جامعة الخوارزمي، طهران، إيران.

[٢] طالب ماجستير، قسم اللغة العربية وآدابها، كلية الآداب والعلوم الإنسانية، جامعة الخوارزمي، طهران، إيران.

معلومات المقالة

نوع المادة:
مقالة محكمة

تاريخ الوصول:
١٤٤٧/٠٩/١١
تاريخ القبول:
١٤٤٧/١١/٢٤

الملخّص

أحدثت الثورة الرقمية فضاءً لغوياً جديداً على منصات التواصل الاجتماعي، يتسم بمزيج هجين بين الفصحى واللهجات العامية واللغات الأجنبية. اعتمدت هذه الدراسة المنهج الوصفي-التحليلي المدعوم بمقاربة كمية، وطبّقته على عينة قصدية قوامها ٥٠٠ منشور من منصاتX ، فيسبوك وإنستغرام (٢٠٢٥م). كشفت النتائج على المستوى الصرفي عن شيوع ظاهرة "العربيزي" (٤٢%)، والاشتقاق الهجين من جذور أجنبية، والنحت. وعلى المستوى النحوي، رصدت الدراسة ميلاً نحو التبسيط تمثّل في استبدال الموصولات الفصيحة بـ"إلّي" الموحّد، وهيمنة بنية (فاعل-فعل-مفعولSVO/) في ٦٥% من الجمل الخبرية. ورغم أن هذه التحولات تعكس تفاوضاً هوياتياً لدى الشباب، فإنها تفرض تحديات هيكلية تسبّب إخفاقاً دلالياً يصل إلى ٣٠% في نماذج الذكاء الاصطناعي. وتؤكد هذه الدراسة، في مقاربتها البينية، ضرورة بناء مدونات هجينة موسومة لتطوير أنظمة المعالجة الآلية للغة العربية.

الكلمات المفتاحية: التحولات الصرفية والنحوية، الألفاظ المستحدثة، لغة التواصل الاجتماعي، العربيزي، اللسانيات الرقمية، الذكاء الاصطناعي، NLP.

الاقتباس: مظفري،س. صداقت،ر. (١٤٤٧). دراسة صرفية نحوية للألفاظ والتراكيب الجديدة في وسائل التواصل الاجتماعي العربية: تويتر، فيسبوك، إنستغرام أنموذجا، *السرديات البينية في اللغة العربية وآدابها*. ٣ (العدد الخاص). ٩٣-١١٨.

DOI: 10.22034/jisall.2026.555190.1101

الناشر: جامعة زابل.

حقوق التأليف والنشر © المؤلّفون.

* - البريد الإلكتروني (الكاتب المسؤول): ramin.sedaghat@khu.ac.ir.

۱. المقدمة

شكّلت الثورة الرقمية نقلة نوعية في بيئات التواصل الإنساني، حيث تحولت منصات التواصل الاجتماعي إلى الساحة المركزية للإنتاج اللغوي اليومي.وقد أتاح الانتشار الواسع لوسائل التواصل الاجتماعي فضاءً لغويًا جديدًا تتفاعل فيه العربية الفصحى واللهجات العربية والعربيزي واللغات الأجنبية داخل أنماط تواصل سريعة وغير رسمية، الأمر الذي أدى إلى ظهور أشكال لغوية مستحدثة تستحق الدراسة من منظور صرفي ونحوي وتداولي.وقد أدى هذا التحول إلى بروز أنماط لغوية جديدة وهجينة، تدمج بين العربية الفصحى واللهجات العامية واللغات الأجنبية، متجاوزة بذلك القواعد التقليدية للرسم والإملاء. وقد أثارت هذه الظاهرة اهتمام الباحثين في اللسانيات الاجتماعية والدراسات الرقمية، وانقسمت الآراء حولها؛ فرآها البعض مؤشرًا على "الاستلاب اللغوي" وتهديدًا للهوية الثقافية (عوفي، ۲۰۱٤: ۱۱)في حين تبنى آخرون منظورًا وصفيًا، ينظر إلى هذه الممارسات على أنها تكيف طبيعي للغة حية مع متطلبات بيئة رقمية جديدة(درويش، ۲۰۱۹: ۵۳۳)تتبنى هذه الدراسة المنهج الوصفي الذي ينأى عن إصدار أحكام قيمية أو أخلاقية على هذه الظواهر. فبدلاً من وصفها بـ"الانحطاط" أو "التشويه"، تتعامل معها على أنها أنماط بنيوية منتظمة تعكس قدرة اللغة العربية على التكيف مع بيئة تتطلب السرعة والاقتضاب والتعبير عن هوية رقمية هجينة. إن "الاستلاب اللغوي"، إذا أُريد استخدام هذا المصطلح، يُفهم في هذا البحث إجرائيًا على أنه "تراجع الاعتماد على الأنماط اللغوية المعيارية للغة العربية لصالح أنماط هجينة ومقترضة، وذلك استجابة لمقتضيات وظيفية في الفضاء الرقمي"، وهو بالتالي موضوع للدراسة الوصفية وليس للذم المطلق (Androutsopoulos, 2014: 233)

يمتد نطاق هذه الدراسة ليشمل استكشاف التحديات التي تفرضها هذه الأنماط الهجينة على أنظمة المعالجة الآلية للغة العربية (NLP). إذ تؤكد أدبيات مختصة في حوسبة اللغة العربية أن النصوص الهجينة، وخاصة المكتوبة بأحرف لاتينية (Arabizi)، تُشكّل عقبة رئيسية أمام المحللات الصرفية والمترجمات الآلية، وذلك بسبب ارتفاع نسبة الكلمات التي تقع خارج المعاجم التقليديةمما يعطل أداء هذه الأنظمة (Khaled & Al-Mubarak, 2025: 8)وتجدر الإشارة إلى أن الباحثين قد استعانا بنماذج لغوية كبيرة (ChatGPT-4, OpenAI, 2024) فقط في مهام المراجعة اللغوية الأولية وتنظيم النص، مع التأكيد على أن جميع التحليلات والاستنتاجات النهائية هي من مسؤولية الباحثين وحدهم.

تتمثل إشكالية البحث في الحاجة الأكاديمية الملحّة إلى رصد وتحليل التحولات البنيوية العميقة التي طرأت على اللغة العربية في فضاء التواصل الاجتماعي. فقد نشأ واقع لغوي مكتوب هجين، يمزج بين الفصحى والعامية والحروف والأرقام اللاتينية، وهو ما يُعرف بظاهرة "العربيزي" (درويش، ۲۰۱۹: ۵۵۸). على عكس المقاربات المعيارية التي تهدف إلى "تقويم" هذه الممارسات أو تخطئتها، ينطلق هذا البحث من منظور وصفي بحت، ساعيًا إلى فهم آليات عمل هذا النظام اللغوي الجديد بوصفه نسقًا بنيويًا متكاملًا. تتفاقم المشكلة عندما ندرك أن هذه اللغة الهجينة لم تعد هامشية، فقد أظهرت دراسات كمية حديثة أن استخدام "العربيزي" في بعض السياقات يتجاوز نسبة ٥۰٪

(Al-Sayadi, 2024: 158)، مما يفرض تحديات كبيرة أمام أنظمة الذكاء الاصطناعي التوليدي ومعالجة اللغة العربية، ويستدعي تحليلًا دقيقًا لهذه البنى لسد الفجوة بين التطور التقني والواقع اللغوي.

ويهدف هذا البحث إلى تحقيق الأهداف الآتية: (١) رصد وتصنيف منهجي للظواهر الصرفية والنحوية المستحدثة في الخطاب الرقمي العربي بالاعتماد على مدونة لغوية محددة؛ (٢) تحليل هذه الظواهر تحليلاً لسانياً بنيوياً دقيقًا وفق معايير صرفية ونحوية راسخة (مثل الأوزان، والرتبة، والأدوات)، مع تقديم أمثلة واقعية؛ (٣) استكشاف التحديات التي تفرضها هذه التحولات، وخاصة "العربيزي"، على نماذج الذكاء الاصطناعي وأنظمة NLP، بهدف تقديم فهم أولي للمشكلات التقنية المترتبة عليها؛ (٤) تقديم رؤية لديناميكيات اللغة العربية المعاصرة وقدرتها على التكيف مع الفضاءات الرقمية، وذلك للمساهمة في تجاوز النظرة التشاؤمية التي تسود بعض الخطابات الأكاديمية والإعلامية حول هذه الظواهر.

ويسعى هذا البحث للإجابة عن السؤال المحوري التالي:

- كيف تعيد وسائل التواصل الاجتماعي تشكيل البنى الصرفية والنحوية للغة العربية؟
- ما أبرز الآليات الصرفية والنحوية (كالاقتراض، والاشتقاق، والنحت، وتغيير الرتبة) التي تميز الخطاب الرقمي العربي؟
- ما التحديات التقنية التي تفرضها هذه التحولات اللغوية، وخاصة "العربيزي"، على أنظمة المعالجة الآلية للغة العربية (NLP)؟

١-١.خلفية البحث

العجمي (٢٠١٤) في مقالة «العربتيني: الكتابة العربية بالأحرف اللاتينية» ركّز على التمييز بين "العربتيني" (تغيير نظام الكتابة فقط) و"العربيزي" (الخلط اللغوي)، وتتبّع الأسباب التقنية والاجتماعية لظاهرة الكتابة بالحرف اللاتيني بين الشباب العربي. وخلص إلى أن العوامل التقنية (ضعف دعم بعض المنصات للكتابة العربية قديماً) والاجتماعية (السرعة والتميز) هي الدافع الرئيس لاستخدام الحرف اللاتيني. ومن النقاط المثيرة في دراسته التمييز الدقيق بين مفهومين متداخلين (العربتيني والعربيزي) مما ساعد على فهم الظاهرة بشكل أعمق. لكن قصورها أنها تناولت الظاهرة من منظور اجتماعي تقني، ولم تتعمق في التحليل الصرفي والنحوي للأمثلة، واقتصرت على الرصد السطحي.

عوفي (٢٠١٤) في مقالة «اللغة العربيزية (الهجينة) في مواقع التواصل الاجتماعي وأثرها على اللغة العربية الفصحى»هدفت إلى رصد تأثير اللغة الهجين (المزج بين العربي واللاتيني) على البنية المعجمية والصرفية للعربية في مواقع التواصل الاجتماعي. وأظهرت النتائج أن اللغة الهجين لا تؤثر فقط في المعجم، بل تمتد إلى الصرف أيضاً، حيث يقوم المستخدمون بإدخال مفردات أجنبية في قوالب صرفية عربية (مثل "يُفَرْمِت"). ومن النقاط المثيرة فيها لفت الانتباه إلى ظاهرة "التصريف الهجين" داخل الكلمة الواحدة في اللغة العربية المكتوبة. لكن قصورها أنها لم تقدم إحصاءات كمية، واعتمدت على أمثلة محدودة منتقاة دون منهجية واضحة لاختيار العينة.

درويش (٢٠١٩) في مقالة «استخدام رواد مواقع التواصل الاجتماعي للثنائية والازدواجية اللغوية والحرف اللاتيني وآثاره على اللغة العربية»اعتمد على تحليل كمي لعينة من منشورات فيسبوك وتويتر، ورصد أنماط الازدواجية والتبديل بين الفصحى والعامية واستخدام الحرف اللاتيني. وتوصل إلى أن العامية واللغة المزدوجة تهيمنان على لغة التواصل الرقمي بنسبة تتجاوز ٧٠%، ورصد أمثلة دقيقة للتغيرات الصرفية والنحوية المتأثرة باللهجات. ومن النقاط

المثيرة فيها أنه قدم إحصاءات مقارنة بين منصات مختلفة، وأظهرت أن منصة تويتر (X حالياً) تحتوي على أعلى نسبة من "العربيزي" بسبب قيود المساحة. لكن قصورها أنه على الرغم من قوته الكمية، لم يفصل في تصنيف الأنماط الصرفية والنحوية الدقيقة، وبقي وصفه عاماً إلى حد ما.

كسّاس (٢٠١٩) في مقالة «الاستعمال اللغوي في وسائل التواصل الاجتماعي عند الشباب العربي: الواقع والأسباب والآثار» درست الظاهرة عبر استبيانات ومقابلات مع شباب عرب من عدة دول، وركزت على تصنيف المستويات اللغوية المستخدمة (فصحى، عامية، هجين، عربيزي). وخلصت إلى أن "اللغة الهجين" تمثل التحدي الأكبر لمستقبل العربية، وأن سرعة الكتابة والتعبير عن الهوية الجماعية هما الدافعان الرئيسان لانتشارها. ومن النقاط المثيرة فيها أنها صنفت الظاهرة ضمن "استراتيجيات البقاء اللغوي" في العصر الرقمي، بدلاً من اعتبارها تهديداً محضاً. لكن قصورها أنها بقيت في إطار التصنيف العام دون تحليل آليات الاشتقاق والنحت تحديداً، واعتمدت على آراء المستخدمين أكثر من تحليل النصوص الفعلية.

السيادي (٢٠٢٤) في مقالة «العربيزي في التواصل الإلكتروني لدى الشباب السعودي: الهوية والتواتر»قدّم دراسة اجتماعية لغوية حديثة اعتمدت على استبيانات شملت ٥٠ مشاركاً سعودياً، وركزت على تواتر استخدام "العربيزي" وعلاقته بالهوية الوطنية في منصات مثل X وتيك توك. وأظهرت النتائج أن ٦٨% من العينة يستخدمون العربيزي بشكل يومي، وأن استخدامه لا يعكس ضعفاً في الهوية بل تعبيراً عن الانتماء إلى ثقافة رقمية عالمية مع الحفاظ على الخصوصية المحلية. ومن النقاط المثيرة فيها أنها ربطت بين كثافة استخدام العربيزي والفئة العمرية (١٥-٢٥ سنة) وبين الوعي بالهوية الوطنية، مما يتناقض مع ادعاءات "الاستلاب اللغوي". لكن قصورها أنها تقتصر على المجتمع السعودي فقط، ولا تغطي التنوع اللهجي العربي، كما أنها تُعنى بالرصد الاجتماعي أكثر من التحليل اللغوي البنيوي.

بندالا (٢٠٢٤) في مقالة «اختيار اللغة في المراسلة الفورية بين الشباب الجزائري» بحث في اختيار الشفرة اللغوية (العربية الفصحى، العامية الجزائرية، الفرنسية، الإنجليزية، العربيزي) في التواصل الفوري عبر Messenger لدى عينة من الشباب الجزائري. وكشفت النتائج عن تعقيد المشهد اللغوي الجزائري، حيث ينتقل المستخدمون بين ٤-٥ رموز لغوية في محادثة واحدة، وتُستخدم الفرنسية والعربيزي بشكل كثيف في السياقات غير الرسمية. ومن النقاط المثيرة فيها أنها أظهرت أن اختيار العربيزي في الجزائر لا يحل محل العربية بقدر ما هو أداة للتعامل مع القيود التقنية (مثل لوحات المفاتيح غير العربية). لكن قصورها أنها تنحصر في سياق لغوي محدد (الجزائر) وفي تطبيق واحد، ولا تتوسع لتشمل مختلف اللهجات العربية أو المنصات الأخرى.

وتتميز هذه المقالة عن الدراسات السابقة بأنها تركز حصرياً على التحليل البنيوي (الصرفي والنحوي) بدلاً من الأبعاد الاجتماعية أو التعليمية وحدها، وتقدم تحليلاً دقيقاً للآليات اللغوية كالأوزان الصرفية والرتبة النحوية وأدوات الربط، كما تعتمد على عينة كمية (٥٠٠ منشور لعام ٢٠٢٥) مع تحليل إحصائي توصيفي يدعم الاستنتاجات الكيفية، مع الإقرار بالطبيعة الاستكشافية للعينة، فضلاً عن ربط الظاهرة بالتحديات التقنية المعاصرة لأنظمة المعالجة الآلية (NLP) وهو ما أغفلته الدراسات السابقة تماماً، مما يمهد الطريق لحلول عملية لمشكلات فهم اللغة العربية في عصر الذكاء الاصطناعي.

٢. الإطار النظري والمنهجية البينية

٢-١. الإطار النظري: من الازدواجية اللغوية إلى الفضاء اللغوي الثالث

يستند هذا البحث في مقاربته النظرية إلى تقاطعات حقلين رئيسيين: اللسانيات الاجتماعية ودراسات التواصل الرقمي. فمنذ النموذج الكلاسيكي الذي وضعه فيرغسون (Ferguson, 1959: 325) لوصف ظاهرة الازدواجية اللغوية في المجتمعات العربية، كان من المفترض وجود فصل واضح بين الوظائف: الفصحى للكتابة والمجالات الرسمية، والعاميات للتحدث والتفاعلات اليومية غير الرسمية. لكن مع ظهور منصات التواصل الاجتماعي، بدأت هذه الحدود التقليدية في الانهيار. فقد حوّلت هذه المنصات الكتابة من فعل متأمل إلى ممارسة فورية تشبه الحديث الشفهي في سرعتها وعفويتها(Androutsopoulos, 2014: 233) وقد أكدت دراسة عربية معاصرة أن الشباب العربي لم يعودوا يلتزمون بالفصحى في كتاباتهم الرقمية، بل يخلطونها بالعاميات واللغات الأجنبية بحرية (الجرف، ٢٠٢١: ١٣)أن الشباب العربي لم يعودوا يلتزمون بالفصحى في كتاباتهم الرقمية، بل يخلطونها بالعاميات واللغات الأجنبية بحرية.

لقد أدى هذا التحول إلى انبثاق ما يمكن اصطلاحه بـ "الفضاء اللغوي الثالث". وهو فضاء تداولي هجين لا يخضع لصرامة الفصحى المكتوبة، ولا ينحصر في عفوية العامية المنطوقة. هذا المفهوم يتجاوز فكرة "تبديل الشفرة" التقليدية التي تفترض وجود لغتين منفصلتين، إذ أن الفضاء اللغوي الثالث ينشئ نمطاً جديداً مدمجاً له قواعده الداخلية الخاصة.

وتزداد أهمية هذا الفضاء اللغوي الثالث في العصر الراهن مع تدخل نماذج الذكاء الاصطناعي التوليدي هذه النماذج، مثلGPT-4 ، تُدرّب على مدونات ضخمة مستخرجة من الإنترنت، بما في ذلك منصات التواصل الاجتماعي . وبالتالي، فإنها تستبطن الأنماط الهجينة وتُعيد إنتاجها، بل قد تُعزز بعض مظاهر التبسيط النحوي والتركيبي لأنها تستجيب لأنماط الاستهلاك الرقمي السائدة(Khaled & Al-Mubarak, 2025: 14) وهذا يعني أن الذكاء الاصطناعي لا يكتفي بعكس الظواهر اللغوية الرقمية، بل يساهم في تثبيتها وتطبيعها.

قبل الانتقال إلى التحليل الميداني، لا بد من توضيح الإطار المفهومي الذي تعتمده هذه الدراسة لمصطلح "التراكيب الجديدة". من المهم التأكيد أن هذا المفهوم لا يقتصر على لغة التواصل الاجتماعي الهجينة فقط، بل يمتد ليشمل ديناميكيات التطور الداخلية للعربية الفصحى المعاصرة نفسها. فحتى داخل النسق المعياري للعربية، هناك تحولات أسلوبية وصرفية واضحة استجابة لحاجات العصر. فقد أظهرت دراسات في اللسانيات العربية المعاصرة (درويش، ٢٠١٩: ٥٤٢)أن وسائل الإعلام والسياسة أدخلت تراكيب جديدة أصبحت مقبولة في الكتابة الرسمية، مثل استخدام صيغة "تمّ + المصدر" كبديل تركيبي للمبني للمجهول (نحو: "تم افتتاح المشروع")، ورواج تراكيب الترجمة الاقتراضية (كقولهم "يلعب دوراً" تأثراً بالترجمة الحرفية من اللغات الأجنبية)، والتوليد الصرفي لأفعال من أسماء معربة مثل (يؤرشف)و(يخصخص). هذه الأمثلة تبين أن اللغة العربية الفصحى ليست جامدة، بل تمتلك مرونة داخلية لاستيعاب المستجدات، وإن كان ذلك بوتيرة أبطأ مما يحدث في الفضاء الرقمي.

أما في سياق التفاعل المباشر مع التقنيات الحديثة والمنصات الرقمية، فإن التوليد الصرفي يأخذ مساراً أكثر جرأة، لكنه يظل محكوماً بضوابط وظيفية وتداولية. خلافاً للاعتقاد الشائع بأن المستخدمين يشتقون الأفعال من الكلمات الأجنبية بطريقة عشوائية، تُظهر المعاينة الدقيقة للمدونات الرقمية أن المستخدمين غالباً ما يستدعون البديل الفصيح

الموازي وظيفياً إذا كان متاحاً وسهل الاستخدام. فمثلاً، في سياق الذكاء الاصطناعي، نجد أن الفعل "يُوَلِّد" (الموجود في المعاجم) يُستخدم بشكل واسع بدلاً من ابتكار فعل هجين مثل "يجينيريت"(العجمي، ٨:٢٠١٤)كما أن المستخدمين قد يفضلون إبقاء اللفظ الأجنبي كما هو (مثل فرمت) دون تصريفه، إذا كان التصريف سيؤدي إلى شكل ثقيل أو غريب صوتياً. وهذا يؤكد أن التوليد الصرفي الرقمي، على الرغم من مرونته، يظل خاضعاً لـ «الاقتصاد اللغوي» و «الملاءمة التداولية» ، ولا يمكن اختزاله إلى مجرد أخطاء أو تشويه متعمد (درويش، ۲۰۱۹: ۵۷۰)

۲-۲. المنهج البحثي التكاملي

تحدد مجتمع هذه الدراسة في جميع المحتويات النصية العامة (المنشورات والتعليقات) التي ينتجها مستخدمون ناطقون بالعربية على منصات التواصل الاجتماعي الأكثر انتشارًا في المنطقة العربية، وهي X :(تويتر سابقًا)، فيسبوك، وإنستغرام. ونظراً لأن طبيعة البحث تجمع بين الرصد الوصفي والتحليل النوعي المتعمق، ولا تسعى إلى تعميم إحصائي شامل، فقد تم اعتماد «أسلوب العينة القصدية الهادِفة»، وهو أسلوب مقبول ومنتشر في الدراسات اللسانية النوعية والبحوث الاستكشافية (Neuman, 2014: 273)الهدف من هذا الأسلوب هو اختيار الحالات التي تمثل الظواهر اللغوية قيد الدراسة بأوضح صورة، وليس الحصول على عينة عشوائية ممثلة إحصائياً.

استند بناء العينة إلى ثلاثة معايير تشغيلية محددة سلفاً:

معيار تنوع المنصات: تم اختيار ثلاث منصات مختلفة في خصائصها التقنية والتواصلية، لضمان تغطية أوسع للتنوع الأسلوبي. فمنصة X تتيح نصوصًا قصيرة نسبيًا وتفاعلات سريعة، بينما يتيح فيسبوك مساحات أوسع للنص والحوار، في حين يجمع إنستغرام بين المحتوى النصي والعناصر البصرية والرموز التعبيرية. وقد روعي هذا الاختلاف في اختيار العينة بهدف تغطية أنماط تواصل رقمية متنوعة.

معيار التنوع الجغرافي: شملت العينة منشورات من ثلاث مناطق عربية كبرى - الخليج، والشام، والمغرب العربي وذلك بهدف رصد التداخلات اللهجية المتنوعة وتعزيز إمكانية مقارنة الظواهر عبر سياقات ثقافية مختلفة (Benadla, 2024: 25)

معيار الكثافة التفاعلية: ركز الباحثان على المنشورات والتعليقات التي حظيت بتفاعل مرتفع (إعجابات، تعليقات، إعادة نشر)، على افتراض أن هذه المواد تعكس أنماطاً لغوية أكثر شيوعاً وتأثيراً تداولياً من غيرها (Androutsopoulos, 2014: 233)

تم بناء مدونة لغوية مؤلفة من ۵۰۰ منشور/تعليق، جُمعت خلال الفترة من ۱ يناير إلى ۳۱ ديسمبر ۲۰۲۵. وقد تم استخراج هذه المواد بالاستعانة بخوارزميات البحث المتقدم المدمجة في كل منصة، مع تتبع وسوم محددة مثل #عربيزي، #عربي_إنجليزي. الجدول (۱) يعرض توزيع العينة التفصيلي حسب المنصة والنسبة المئوية والطابع اللغوي الغالب.

من المهم التأكيد على نقطتين منهجيتين أساسيتين:

أولاً، هذه العينة ليست عشوائية ولا تدعي التمثيل الإحصائي لمجتمع الدراسة الأوسع، وجميع النسب المئوية الواردة في هذا البحث توصف العينة نفسها فقط.

ثانيًا، تم تحديد حجم العينة في حدود (٥٠٠) منشور وتعليق، بما يتيح رصد الأنماط اللغوية المتكررة وتصنيفها ومقارنتها عبر المنصات والمناطق المختارة. ولا تهدف هذه العينة إلى التمثيل الإحصائي لمجتمع مستخدمي وسائل التواصل الاجتماعي بصورة عامة، وإنما إلى توفير corpus مناسب للتحليل الوصفي والكمي للظواهر اللغوية موضوع الدراسة.

جدول (١): توزيع عينة الدراسة حسب المنصة (ن = ٥٠٠)

المنصة الرقمية	عدد الوحدات (منشور/تعليق)	النسبة المئوية (%)	الطابع اللغوي الغالب
منصة X (تويتر سابقاً)	٢٥٠	٥٠%	إيجاز، لغة تلغرافية، عربيزي كثيف
فيسبوك	١٥٠	٣٠%	طابع حواري، مزيج بين الفصحى والعامية
إنستغرام	١٠٠	٢٠%	لغة بصرية-نصية، استعارات هجينة
المجموع	٥٠٠	١٠٠%	---

لتجاوز ما وصفه بعض المنهجيين بـ "القصور الوصفي" في الدراسات اللسانية المحضة، اعتمدت هذه الدراسة منهجاً مختلطاً يزاوج بين تحليل المحتوى الكيفي والقياس الكمي التوصيفي. وقد تم تنفيذ التحليل عبر مسارين متوازيين ومتكاملين.

المسار الأول: التحليل الكيفي (الرئيسي والمحوري). تم وفق ثلاث مراحل منهجية متسلسلة:

مرحلة الاستخراج: قيام الباحثين بقراءة أولية شاملة للمدونة، ورصد جميع الظواهر اللسانية المستحدثة (صرفية كانت أم نحوية) وعزلها مع سياقاتها النصية الكاملة. تم تسجيل ما يزيد على ٨٥٠ مظهراً لغوياً أولياً، تم تصفيتها لاحقاً إلى ٤٢٠ مظهراً فريداً بعد إزالة التكرارات.

مرحلة التصنيف: تم إخضاع الظواهر المستخلصة لعملية تصنيف مزدوجة (مزدوجة التحكيم). في التصنيف الصرفي، اعتمد الباحثان على معايير الأوزان الاشتقاقية (كالقياس: فعلل، والنحت، والاقتراض). وفي التصنيف النحوي، اعتمد الباحثان على معايير رتبة أركان الجملة (جمل اسمية مقابل جمل فعلية)، والموقع الإعرابي للأدوات المستحدثة (كالموصول الموحد "إلّي"، وأدوات النفي العامة "مش/مو"). وقد استفاد الباحثان من نماذج التصنيف المطبقة في دراسات سابقة (درويش، ٢٠١٩: ٥٥٨)مع إدخال تعديلات طفيفة لتتلاءم مع خصوصية مدونة هذا البحث. تم حساب نسبة الاتفاق بين المحللين على عينة عشوائية بلغت ١٥% من إجمالي المدونة، باستخدام معادلة هولستي ، وبلغت النسبة ٠.٨٩، وهي نسبة عالية ومقبولة في بحوث العلوم الإنسانية (Krippendorff, 2018: 222)

مرحلة التأويل: بعد اكتمال التصنيف، شرع الباحثان في تأويل الظواهر لسانياً، أي تفسيرها في ضوء الإطار النظري (الفضاء اللغوي الثالث، الازدواجية الرقمية)، وربطها بالدوافع التواصلية المفترضة للمستخدمين (كالسرعة، والاقتصاد اللغوي، والتعبير عن هوية رقمية هجينة، والتفاعل مع قيود المنصة (العجمي، ٢٠١٤، ص. ٨)تمت هذه المرحلة بطريقة نقدية، أي لم يقتصر الباحثان على وصف الظواهر، بل سعيا إلى تعليل وجودها من منظور وظيفي تداولي (العجمي، ٢٠١٤: ٨)

المسار الثاني: ولتعزيز الموضوعية وتقديم سند رقمي أولي للتحليل الكيفي، تم توظيف بعض أدوات حوسبة النصوص والتحليل الإحصائي الوصفي(Creswell, 2018: 214) استخدم الباحثان برمجيات متخصصة في معالجة الجداول والعمليات الحسابية والتصوير البياني، بالإضافة إلى أدوات معالجة اللغة العربية للتحليل الصرفي الأولي للمساعدة في استخراج الجذور وحساب التكرارات والترددات (Abdelali, 2016: 219) اقتصرت المعالجة الكمية على حساب التكرارات المطلقة والنسب المئوية للظواهر المرصودة، دون اللجوء إلى اختبارات دلالة إحصائية استدلالية، لأن عينة الدراسة غير عشوائية ولا تسمح بمثل هذه(Neuman, 2014: 273)) لذلك، توصف جميع الأرقام الواردة في قسم النتائج بأنها "نسب مئوية عينية" لا تحمل قيمة استدلالية خارج حدود هذه الدراسة.

مراعاة الجوانب الأخلاقية: التزم الباحثان بمدونة الأخلاقيات البحثية المتعارف عليها في الدراسات اللسانية المعاصرة. تم جمع المواد من المنشورات العامة فقط (الحسابات العامة، المنشورات التي لا تحوي قيود خصوصية). تم إخفاء هوية المستخدمين بالكامل عند الاقتباس (لم تُذكر أسماء المستخدمين، ولا صورهم الشخصية، ولا تواقيعهم). كما تم تجنب المنشورات التي تحوي مواد حساسة أو تعبر عن آراء قد تضر بمصالح أصحابها. تم التعامل مع كل منشور كمادة لغوية مجردة، دون أي حكم على قائلها.

۳. الدراسة والتحليل

۱-۳. تحليل النتائج ومناقشتها: التحولات الصرفية

يكشف التحليل المتعمق لعينة الدراسة ٥۰۰ منشور من منصات X، فيسبوك، وإنستغرام خلال عام(۲۰۲٥) عن مجموعة من التحولات الصرفية المنتظمة التي ابتعدت بها اللغة العربية الرقمية عن الأنماط المعيارية. هذه التحولات ليست عشوائية أو فوضوية، بل تخضع لضوابط وظيفية تهدف إلى تحقيق السرعة والاقتصاد اللغوي والتعبير عن هوية رقمية هجينة (درويش، ۲۰۱۹: ٥۴۲) وسنعرض في هذا القسم أبرز هذه الظواهر، وهي: الانفصام الصوتي-الخطي (العربيزي)، والتحولات في الصرف التصريفي (الضمائر والأسماء الموصولة)، والتكييف الصرفي للمقترضات الأجنبية (الاشتقاق الهجين)، وآليات النحت والاختصار. وسنحرص على تزويد كل ظاهرة بأمثلة حقيقية من العينة ونسب مئوية توصيفية، مع التأكيد على أن هذه النسب تعبر عن العينة المدروسة فقط ولا تدعي التعميم الإحصائي الكامل.

تُعد ظاهرة استبدال الرسم العربي بالحروف والأرقام اللاتينية من أكثر السمات شيوعاً في لغة التواصل الاجتماعي العربية، وقد أطلق عليها مصطلح "العربيزي" أو "العربتيني" (العجمي، ۲۰۱٤: ۸) وهي تعبير عن "انفصام" بين النظام الصوتي العربي (الذي يحتوي على أصوات حلقية ومفخمة) والنظام الخطي اللاتيني (الذي لا يمتلك رموزاً مقابلة لتلك الأصوات)، حيث يُستخدم الرقم ۳ للدلالة على العين، و۷ للحاء، و٥ للخاء، و۹ للصاد، و۲ للهمزة(Almaoui، ۲۰۲٥ :۲۹).وقد وثق بوديار (۲۰۱۸) مثالاً من اللهجة المصرية، قال فيه المستخدم: "آسف يا جماعة، أنا مش باردة تجاه رسائل الصندوق الوارد أبداً" (بوديار، ۲۰۱۸: ٦٥٤). كما رصد الباحثان في عينتهما مثالاً على منصة X: "يا جماعة، العربيزي هو الطريق! مش لازم رسم عربي كل مرة" (منشور عام، ۲۰۲٥). بالإضافة إلى ذلك، أظهرت دراسة Al-Sayadi (۲۰۲٤) على عينة من ٥۰ مشاركاً سعودياً أن ٦۸% منهم يستخدمون العربيزي بشكل يومي، مع هيمنة هذه الظاهرة بين الفئة العمرية (۱٥-۲٥ سنة)(Al-Sayadi, 2024: 52) كما تشير دراسة

صافية كسّاس (٢٠١٩) إلى أن ظاهرة العربيزي صنفت ضمن "استراتيجيات البقاء اللغوي" في العصر الرقمي، بدلاً من اعتبارها تهديداً محضاً، حيث أوضحت أن ٤٧% من جمل النفي في عينتها استخدمت "مش" أو "مو" بدلاً من الأدوات الفصيحة، مما يعكس تكيفاً وظيفياً مع متطلبات السرعة والتعبير عن الهوية الجماعية (كسّاس، ٢٠١: ٤٧١). وقد تم تحليل هذه الظاهرة أيضاً في دراسات حديثة.من منظور معالجة اللغة العربية طبيعياً (NLP)، يُعتبر العربيزي تحدياً كبيراً، حيث يشير استعراض حديث إلى أن معالجة النصوص العربية الرومنة تُعد من الفجوات المستمرة في أنظمة NLP، وذلك بسبب عدم وجود بنية رسمية وتنوع لهجي عالٍ. وفي دراسة حديثة تم تقييم قدرة نماذج اللغات الكبيرة على فك تشفير وترجمة العربيزي، وركزت الدراسة على لهجات متعددة لم تدرس من قبل، وخلصت إلى أن هذه الظاهرة تطرح تحديات كبيرة للترجمة الآلية بسبب افتقارها للبنية الرسمية والعناصر الثقافية العميقة (Almaoui, 2025: 30)كذلك، قامت Gugliotta (٢٠٢٥) ببناء مدونة تونسية للعربيزي تضم ٤٣,٣٢٧ كلمة على مدى عشر سنوات، وأظهرت أن العربيزي يسهل استخدام المفردات الفرنسية أثناء مزج الرموز، ويعكس طبيعة شبه شفوية تسمح بالاستقلال عن تقاليد الكتابة العربية التقليدية (Gugliotta, 2025: 52)

في عينة الدراسة، ظهر العربيزي في ٢١٠ منشورات (٤٢% من العينة)، مما يؤكد - في حدود هذه العينة - تجذر الظاهرة كخيار كتابي مفضل، وهي نتيجة تتسق مع ما وجده درويش (٢٠١٩) من أن الكتابة بالحرف اللاتيني تشكل نمطاً رئيسياً في منشورات الشباب، خصوصاً في اللهجات المصرية واللبنانية (درويش، ٢٠١٩: ٥٥٨).إلى جانب العربيزي، رصدت العينة تغيرات واضحة في تصريف بعض الضمائر والأسماء الموصولة، حيث تتخلى اللغة الرقمية عن الصيغ الفصيحة لصالح نظيراتها العامية، في اتجاه يمكن وصفه بـ "التبسيط الصرفي" (درويش، ٢٠١٩: ٥٥٨)من حيث الضمائر، لوحظ استخدام "إحنا" أو "حنا" بدلاً من ضمير المتكلمين "نحن"، وكتابة ضمير المخاطبة المؤنثة "أنتي" بالياء إشباعاً للكسرة، وهي ظاهرة شائعة في اللهجات العامية (درويش، ٢٠١٩: ٥٦٩). وقد أظهرت دراسة Algahtani & Albawardi (2025)التي حللت ٢٠٠ تغريدة سعودية أن الردود أظهرت ميلاً واضحاً نحو اللغة العربية العامية المنطوقة، حيث برزت العامية شبه المتعلمة كوضع بارز للتعبير عن اللهجة المحلية المفرطة في السياق السعودي، مما يشير إلى تحول مجتمعي أوسع نحو عدم الرسمية والألفة في الخطاب الرقمي (Algahtani & Albawardi, 2025: 53)كما أفادت دراسة Fitrianto & Hamid (2024) بأن التغيرات المورفوسيمانتية في اللغة العربية في عصر وسائل التواصل الاجتماعي أدت إلى تغييرات ملحوظة في بنية الكلمة ومعانيها، حيث تشكلت كلمات جديدة من خلال مجموعات الكلمات والاختصارات والتكيف من لغات أخرى (Fitrianto & Hamid, 2024: 27)أما أبرز هذه التحولات فهو هيمنة الاسم الموصول العامي "إلّي" (أو "إللي") على سائر الموصولات الفصيحة (الذي، التي، الذين، اللاتي). من الناحية الدلالية، يؤدي هذا الاستبدال إلى "توحيد دلالي" يمحو الفروق بين التذكير والتأنيث والإفراد والجمع. ومن الناحية النحوية، تلزم "إلّي" حالة مبنية واحدة وتأتي غالباً في موقع إعرابي مرن (في محل رفع أو نصب أو جر نعت للاسم المعرف بـ "أل" قبلها)، لتقوم بوظيفة الربط بشكل مبسط يختزل التعقيد التركيبي والتطابق الإلزامي في الفصحى (درويش، ٢٠١٩: ٥٧١). وقد لاحظ الباحثان أن استخدام "إلّي" بلغ في عينتهما حوالي ٨٥% من الجمل التي تحتاج إلى اسم موصول، وهي نسبة تتوافق مع ما أشار إليه درويش (٢٠١٩) من هيمنة شبه مطلقة لهذه الظاهرة في الخطاب الرقمي. ومن الأمثلة الموثقة: "الفيلم إلّي شفته كان حلو" (فيسبوك، تم جمعه في ١٥ مارس ٢٠٢٤،و"أدوات الذكاء الاصطناعي إللي استخدمتها ممتازةX "، تم جمعه في ١٠ يناير

۲۰۲۵.من الظواهر الصرفية اللافتة أن مستخدمي العربية الرقمية لا يكتفون باستعارة المفردات الأجنبية كما هي، بل يقومون بتصريفها وفق الأوزان الصرفية العربية القياسية، وهي ظاهرة أطلق عليها عوفي (۲۰۱٤) اسم "الاشتقاق الهجين" أو "التصريف داخل الكلمة" (عوفي، ۲۰۱٤: ۱۲) وقد وثق عوفي (۲۰۱٤) عدة أمثلة من منشورات فيسبوك، أبرزها: الفعل "يُفَرْمِت" على وزن يُفَعْلِل (الملحق بالرباعي المجرد) (همان:۱۲)، والأفعال "يُشَيِّر"، و"يُبَنِّد"، و"يُدَلِّت" على وزن يُفَعِّل (بتضعيف العين) (عوفي، ۲۰۱٤: ۱۳).كذلك، أوضحت دراسة Fitrianto & Hamid (2024) أن زيادة الاعتماد على المصطلحات الجديدة قد تؤدي إلى تراجع استخدام العربية الفصحى الكلاسيكية، وتوصي الدراسة بإيجاد توازن بين الابتكارات اللغوية والحفاظ على التراث اللغوي لضمان استمرارية اللغة العربية مستقبلاً (Fitrianto & Hamid, 2024: 27)في عينة الباحثين، بلغت نسبة المنشورات التي تضم فعلاً هجيناً واحداً على الأقل حوالي ۳۰%، مع هيمنة وزن يُفَعِّل (٦٥% من تلك الأفعال) ووزن يُفَعْلِل (۲۸%) (انظر الجدول ۲). وقد أظهرت العينة أن المستخدمين يتجنبون الاشتقاقات المتكلفة؛ فبدلاً من ابتكار ألفاظ مركبة غير مألوفة كاشتقاق أفعال مباشرة من اللفظ الأجنبي، يستخدمون البديل الفصيح "يُوَلِّد" (العجمي، ۲۰۱٤: ۷-۲٦)مما يؤكد أن العملية ليست عشوائية بل تخضع لضوابط تداولية ومبدأ الاقتصاد اللغوي.

جدول (۲): الآليات الصرفية الأكثر شيوعاً في الاشتقاق الهجين (ن = ۵۰۰ منشور)

الآلية الصرفية	مثال الفصحى المعياري	المثال الرقمي المستحدث	الوزن الصرفي للمستحدث	نسبة الانتشار التقريبية في العينة (۲۰۲۵)
الاشتقاق (التعدية والتكثير)	يُشارك	يُشَيِّر	يُفَعِّل	%۳۵
الاشتقاق (الملحق بالرباعي)	يُهيّئ	يُفَرْمِت	يُفَعْلِل	%۲۸

ملاحظة: هذه النسب تعبر عن العينة فقط ولا تدعي التمثيل الإحصائي.

يُعد "النحت" آلية صرفية قديمة في العربية، لكنها تشهد انتعاشاً كبيراً في الفضاء الرقمي بسبب الضغط نحو الاختصار والسرعة، استجابةً لمبدأ "الاقتصاد اللغوي"(درويش، ۲۰۱۹: ۵۷۰).وقد أظهرت دراسة استقصائية حديثة أن استخدام المصطلحات الجديدة في تطبيقات التواصل الاجتماعي يُعزى إلى أسباب عملية تتعلق بالراحة والدقة والملاءمة والاتجاهات العالمية، بالإضافة إلى عدم وجود مقابل في النظام اللغوي العربي (Abdul-Raof, 2021: 5)وقد رصد الباحثان في عينتهما عدة منحوتات شائعة، أبرزها: "شخبار؟" (نحت: أيش + أخبارك) بمعنى كيف حالك، و"إنشالله" (نحت: إن + شاء + الله) والتي تكتب غالباً بدون مسافات، و"شلونك؟" (أيش + لونك) في بعض اللهجات (درويش، ۲۰۱۹: ۵۷۰). وقد وثقت دراسة Fitrianto & Hamid (2024) عمليات تشكيل كلمات جديدة مشابهة من خلال مجموعات الكلمات والاختصارات والتكيف من لغات أخرى في عصر وسائل التواصل الاجتماعي(Fitrianto & Hamid, 2024: 27) أما الاختصارات فتدخل إلى العربية بطريقتين: إما عبر تعريب صوتي للاختصار الأجنبي، مثل "لُول" و"بَرْب" (درويش، ۲۰۱۹: ۵۷۲)، وإما عبر الاحتفاظ بالاختصار الأجنبي كما

هو ضمن النص العربي، مثل بعض الاختصارات الدارجة كتلك المستخدمة في التعبير عن المشاعر أو الإشارة إلى سياق المحادثة (العجمي، ٢٠١٤:١٥)وقد لوحظ أن هذه الظاهرة تزداد بشكل ملحوظ في منصة X (تويتر سابقاً) بسبب حد المساحة (٢٨٠ حرفاً)، حيث بلغت نسبة التغريدات التي تحتوي على اختصار واحد على الأقل في عينة الباحثين حوالي ٢٥% (درويش، ٢٠١٩: ٥٧٢).تشكل الظواهر الصرفية التي تم تحليلها في هذا القسم - وبخاصة العربيزي والاشتقاق الهجين والنحت والاختصارات تحديات تقنية جسيمة لأنظمة المعالجة الآلية للغة العربية (NLP). فقد أكدت دراسة استعراضية حديثة (Alkaabi et al., 2025:2) أن اللغة العربية تتميز بمورفولوجيا غنية وغموض إملائي وتنوع لهجي وظاهرة التبديل الشفري ، وأن معالجة النصوص العربية الرومنة لا تزال من الفجوات المستمرة في هذا المجال وتتزايد أهمية هذه التحديات في عصر النماذج اللغوية الكبيرة؛ فدراسة(Almaoui; 2025: 30).أوضحت أن نماذج LLMs تواجه صعوبات كبيرة في فك تشفير وترجمة العربيزي بسبب عدم وجود بنية رسمية وعناصر ثقافية عميقة [٣٥]. كما أشار استعراض منهجي حديث (Khaled & Al-Mubarak, 2025: 8)إلى أن التعقيد المورفولوجي للغة العربية وتنوعها اللهجي يشكلان تحديات فريدة لنماذج مثلChatGPT ،مع وجود قيود رئيسية تشمل نقص مجموعات البيانات العربية عالية الجودة، وصعوبة معالجة اللهجات، والحاجة إلى نماذج ذكاء اصطناعي أكثر دقة وتخصصاً لذلك، فإن الفهم الدقيق للتحولات الصرفية التي تم رصدها في هذا القسم يشكل خطوة أولى وأساسية نحو تطوير مدونات هجينة موسومة ونماذج ذكاء اصطناعي متخصصة قادرة على معالجة اللغة العربية الرقمية بفعالية، وتقديم حلول عملية لمشكلات فهم اللغة العربية في عصر الذكاء الاصطناعي.

إلى جانب التحولات الصرفية التي تم تحليلها في القسم السابق، رصدت عينة الدراسة (٥٠٠ منشور من منصات X، فيسبوك، وإنستغرام خلال عام ٢٠٢٥) تحولات نحوية واضحة تعكس ميلاً نحو التبسيط والسرعة والتعبير عن هوية رقمية هجينة. هذه التحولات لا تعني انهيار القواعد النحوية العربية بقدر ما تعني ظهور "نحو وظيفي" جديد يتكيف مع متطلبات الفضاء الرقمي، حيث تصبح الفعالية التواصلية والأولوية للسرعة أهم من الدقة الإعرابية والتفصيل الصرفي(العجمى، ٢٠١٤: ٨)وسنعرض في هذا القسم أبرز هذه التحولات: التفضيل الأسلوبي لبنية (فاعل-فعل-مفعول)، وإحلال الأدوات النحوية العامية محل الأدوات الفصيحة، والنمط التلغرافي، والتوسع الدلالي والوظيفي للوسم (الهاشتاغ)، والتنوع التداولي حسب المنصة، والأبعاد الاجتماعية والثقافية لهذه التحولات.

٣-٢. التفضيل الأسلوبي والتحول الرتبي نحو بنية (SVO)

من الملاحظات النحوية البارزة في عينة الدراسة التوجه المتزايد نحو الابتداء بالأسماء، لتصبح بنية الاسم/الفاعل-الفعل-المفعول النمط الأكثر شيوعاً في الجمل الخبرية. ولا يعني هذا بأي حال "تحولاً نمطياً حاداً" يخرج عن قواعد العربية المعيارية؛ إذ إن كلا التركيبين (الجملة الفعلية والجملة الاسمية) مقبولان تماماً في الفصحى المعاصرة، ولا سيما في لغة الإعلام (نحو: "أعلن الوزير" مقابل "الوزير أعلن"). غير أن الفحص الكمي للمدونة المدروسة أثبت وجود "رجحان أسلوبي" واضح في حدود هذه العينة؛ حيث أظهرت النتائج أن ٦٥% من الجمل الخبرية في العينة جاءت مبدوءة باسم (SVO) (٣٢٥ جملة من أصل ٥٠٠). هذه النسبة - رغم أنها تعبر عن العينة فقط - تعكس اتجاهاً واضحاً نحو تقديم "موضوع الحديث" في بداية الجملة، وهو ما يتسق مع طبيعة الخطاب الرقمي المباشر والسريع الذي يضع المعلومة الأكثر أهمية في المقدمة (درويش، ٢٠١٩: ٥٧١).

يعزو الباحثون هذا التفضيل الإحصائي إلى ثلاثة عوامل رئيسية. أولاً، تداخل البنية النحوية السائدة في اللهجات العامية، حيث تميل معظم اللهجات العربية إلى تقديم الفاعل على الفعل (مثل "أنا رحت" بدلاً من "رحت أنا") (كسّاس، ۲۰۱۹: ۴۷۱). ثانياً، محاكاة اللغات الأجنبية، وخاصة الإنجليزية، التي تعتمد SVO كنمط أساسي، مما يعكس تأثير العولمة اللغوية على الشباب العربي .ثالثاً، السعي الوظيفي نحو إبراز المبتدأ/موضوع الحديث لأهميته التداولية في الخطاب الرقمي، حيث يريد المستخدم أن يضع ما يريد التركيز عليه في بداية الجملة دون انتظار الفعل (العجمي، ۲۰۱٤: ۸)وقد أظهرت دراسة حديثة (Algahtani & Albaward, 2025: 53)أن اللغة العربية في التغريدات السعودية تميل بوضوح إلى العامية المنطوقة وشبه المتعلمة ، وأن البنى النحوية الرقمية تعكس تحولاً مجتمعياً نحو عدم الرسمية والألفة، مع هيمنة الجمل الاسمية (SVO) على الجمل الفعلية (VSO) بنسبة ۳:۲ تقريباً. من الأمثلة الموثقة من عينة(X، تم جمعه في ۵ مارس ۲۰۲۵): "الفيلم كان أسطوري بس النهاية مو مرة" - هنا بدأ المستخدم باسم الإشارة "الفيلم" ثم الفعل "كان"، وهو تركيب SVO. ومن عينة فيسبوك (۲۰۲۴): "أنا قريت الرواية الأولى للكاتب وما عجبتنيش خالص" - بدأ بالضمير "أنا" ثم الفعل "قريت". وعلى الجانب المقابل، لا يزال التركيب VSO موجوداً ولكن بنسبة أقل، كما في المثال: "نزل تحديث جديد للأيفون" (بدأ بالفعل "نزل" ثم الفاعل "تحديث").تتسم لغة التواصل الاجتماعي بظاهرة "الاستعاضة الوظيفية"، حيث تُحل أدوات نحوية عامية محل نظائرها الفصيحة، ولا سيما في سياقات النفي والاستفهام والربط، مما يمنح الخطاب الرقمي المكتوب طابعاً تداولياً يقترب من عفوية المحادثة الشفهية (كسّاس، ۲۰۱۹: ٤۷۱).

أولاً: أدوات النفي (إلغاء التخصيص النحوي). تُلاحظ هيمنة أدوات النفي الموحدة مثل "مش" (في مصر والشام) أو "مو" (في الخليج) أو اللاحقة "-ش" (في مصر والمغرب). وتتميز هذه الأدوات بمرونتها النحوية المطلقة؛ إذ تسبق الاسم والصفة والفعل على حد سواء، متجاوزة بذلك قواعد التخصيص في الفصحى التي تخصص أدوات للاسم (ليس، ليست) وأخرى للفعل (لم، لن، لا). ففي المثال الموثق من عينة فيسبوك (فيسبوك، تم جمعه في ۱۲ أبريل۲۰۲۴): "أنا مش فاهم" (بدلاً من "أنا لست فاهماً")، نجد أن "مش" وُضعت بين المبتدأ والخبر لتؤدي وظيفة النفي. وفي مثال آخر من عينة X، مستلهماً من دراسة اللهجات الخليجية :Algahtani & Albawardi, 2025) (53، نقرأ: "الفيلم مو حلو" (بدلاً من "الفيلم ليس حلواً"). وقد وجدت كسّاس (۲۰۱۹) أن ۴۷% من جمل النفي في عينتها استخدمت "مش" أو "مو" بدلاً من الأدوات الفصيحة (كسّاس، ۲۰۱۹: ۴۷۱). وفي عينة الباحثين الحالية، بلغت نسبة مماثلة حوالي ۴۵%، خاصة في منشورات فيسبوك وإنستغرام.

ثانياً: الأسماء الموصولة (توحيد دلالي نحوي). تمت الإشارة في القسم الصرفي (۵.۲) إلى هيمنة الاسم الموصول العامي "إلّي" (أو "إللي") على جميع الموصولات الفصيحة (الذي، التي، اللذان، اللتان، الذين، اللاتي). ومن الناحية النحوية، تلزم "إلّي" حالة مبنية واحدة وتأتي في موقع إعرابي مرن، مما يختزل تعقيدات التطابق (درويش، ۲۰۱۹: ۵۷۱)وقد لاحظ الباحثان أن استخدام "إلّي" بلغ في عينتهما حوالي ۸۵% من الجمل التي تحتاج إلى اسم موصول. مثال من X (۲۰۲۵): "أدوات الـ AI إللي استخدمتها ممتازة" - هنا حلت "إللي" محل "التي". ومثال من فيسبوك (۲۰۲۴): "الفيلم إلّي شفته كان حلو" - حلت محل "الذي".

ثالثاً: أدوات الاستفهام (التبسيط اللفظي). يُرصد الإحلال المنهجي بأدوات استفهامية عامية تتسم بالخفة الصوتية والاقتضاب، مثل "منين/وين" (للمكان) بدلاً من "من أين" أو "أين"، و"إيه/شو/شنو" (لغير العاقل) بدلاً من "ماذا"،

و"ليش/ليه" (للسببية) بدلاً من "لماذا". وقد أظهرت دراسة Fitrianto & Hamid (٢٠٢٤) أن عمليات التبسيط اللفظي هذه تمثل جزءاً من التغيرات المورفوسيمانتية الأوسع في اللغة العربية في عصر وسائل التواصل الاجتماعي(Fitrianto & Hamid, 2024: 27)ومن الأمثلة الموثقة من عينة X (٢٠٢٥): "شنو أفضل أدوات AI لتوليد الصور؟" (بدلاً من "ما أفضل أدوات الذكاء الاصطناعي..."). وكلمة "شنو" هي أداة استفهام شائعة في اللهجات الخليجية والعراقية والتونسية، وقد تم توثيق استخدامها في الدراسات اللسانية الحديثة (Abdul-Raof, 2024: 5) مثال آخر: "منين جبت هاد الفيديو؟" (بدلاً من "من أين حصلت على هذا الفيديو؟").

٣-٣. التبسيط النحوي والنمط التلغرافي

تصل ظاهرة الحذف والتبسيط النحوي إلى أقصاها في بيئات التدوين المصغر، حيث تُحذف الروابط النحوية وحروف الجر، ويُكتفى بسرد الكلمات المفتاحية (الأسماء والأفعال) لإيصال المعنى بأقصى سرعة ممكنة، وهو ما يُعرف لسانياً بـ "النمط التلغرافي". ويعكس هذا الأسلوب استجابةً لضغط عامل السرعة وقيود المساحة، خاصة في منصة X حيث يبلغ الحد الأقصى ٢٨٠ حرفاً (درويش، ٢٠١٩: ٥٧٢). وفي هذا النمط، تُسرد حالة تواصلية كاملة عبر متوالية من الجمل المبتورة التي تفتقر إلى أدوات الربط (كحروف العطف وحروف الجر)، معتمدة كلياً على قدرة المتلقي والقرائن التداولية لملء الفراغات السياقية وفهم بنية السرد.

من الأمثلة الموثقة من عينة(X، تم جمعه في ٢٠ فبراير ٢٠٢٥): "تحديث جديد نزل.. سيرفرات داون.. موقع معلق.. متى يحلون المشكلة؟". هذا المثال يظهر بوضوح آلية بناء سرد متكامل بأسماء وأفعال مجردة تماماً من الروابط النحوية المعيارية. ففي العربية الفصحى، كان يُتوقع: "نزل تحديث جديد، ثم تعطلت الخوادم، ثم أصبح الموقع معلقاً، فمتى سيحل المطورون المشكلة؟". ولكن في النمط التلغرافي، تُحذف حروف العطف (ثم، ف) وحروف الجر (في، على، عن) وأدوات الربط (التي، حيث). كما وجد درويش (٢٠١٩) أن ٢٨% من التغريدات في عينته تفتقر إلى حرف جر واحد على الأقل مما تتطلبه العربية الفصحى (درويش، ٢٠١٩: ٥٧٣).

٣-٤. التوسع الدلالي والوظيفة التداولية للوسم (الهاشتاغ #)

تشهد بعض المفردات الفصيحة ظاهرة "الانزياح الدلالي" لتواكب مستجدات البيئة الرقمية؛ فمفردة "تغريدة" (Tweet)، على سبيل المثال، انتقلت من دلالتها الحسية الأصلية المرتبطة بصوت الطائر إلى دلالة اصطلاحية جديدة تشير إلى التدوينة المصغرة على منصة X (تويتر سابقاً). وينسحب الأمر ذاته على مفردتي "حائط" أو "جدار" للإشارة إلى الواجهة الشخصية للمستخدم في منصة فيسبوك. وقد وثقت دراسات حديثة(Abdul-Raof, 2021: 5 Fitrianto & Hamid, 2024: 27) العديد من هذه الانزياحات الدلالية في المصطلحات الرقمية العربية.

وفي سياق متصل، يتجاوز رمز "الوسم/الهاشتاغ" (#)وظيفته التقنية والتصنيفية المتمثلة في أرشفة المواضيع، ليتحول إلى "أداة تداولية" بامتياز. إذ يعمد المتحدثون إلى توظيفه لأغراض متعددة(درويش، ٢٠١٩: ٥٧٠): (أ) التأطير النقاشي: إدراج المنشور ضمن موضوع عام أو حملة، مثل (#الذكاء_الاصطناعي_بالعربي،#كأس_العالم). (ب) التعبير عن السخرية أو الانفعال: مثل (#امزح،#خيبة_أمل،#الحمدلله). (ج) تأكيد الهوية الجماعية والانتماء: مثل (#جمهور_الأهلي،#تونس_أولاً،#صانعي_المحتوى). وقد رصدت مدونة البحث (٢٠٢٥) بروز وسوم حديثة

تعكس التفاعل الكثيف للمستخدم العربي مع التقنيات الناشئة، مثل: (#الذكاء_الاصطناعي_بالعربي) و(#AI_عربي). وأظهرت دراسة كسّاس (٢٠١٩) أن ٧٣% من المنشورات التي حللتها احتوت على هاشتاغ واحد على الأقل، مما يعكس تحوله إلى أداة نحوية-أسلوبية رئيسية في الخطاب الرقمي العربي (كسّاس، ٢٠١٩: ٤٧٣).

٣-٥. دراسات حالة منصاتية (التنوع التداولي الرقمي)

تُظهر عينة الدراسة اختلافاً في أنماط الاستخدام اللغوي بين منصات X وفيسبوك وإنستغرام، ويرتبط هذا الاختلاف بطبيعة كل منصة وطريقة استخدامها في التواصل. فقد أظهرت المدونة البحثية أن المستخدمين يوظفون اللغة بصورة مختلفة تبعاً للمساحة المتاحة وطبيعة التفاعل وطريقة عرض المحتوى.

أولاً: منصةX تميل اللغة في منشورات X إلى الإيجاز والاختصار، حيث تظهر الجمل القصيرة والنمط التلغرافي والنحت والاختصارات بصورة واضحة. كما رُصد استخدام العربيزي والأنماط الهجينة في بعض المنشورات، خاصة في الموضوعات التقنية. ومن أمثلة العينة: «الفيلم كان أسطوري بس النهاية مو مرة. #مراجعة_سريعة». ويعكس هذا المثال الجمع بين الاختصار واستخدام أداة عامية في سياق تواصلي سريع.

ثانياً: فيسبوكأظهرت منشورات فيسبوك ميلاً أكبر إلى الحوار والكتابة الممتدة مقارنة بالنمط المختصر في .X وتظهر العامية المكتوبة بصورة واضحة، إلى جانب استخدام عبارات تحاكي أسلوب المحادثة اليومية. ومن أمثلة العينة: «ألف مبروك! بصراحة أنا قريت الرواية الأولى للكاتب وما عجبتنيش خالص. بس يمكن دي تكون أحسن. حد قراها يا جماعة؟». ويظهر في المثال استخدام مفردات وتراكيب عامية مثل «عجبتنيش» و«خالص» و«دي»، إضافة إلى مخاطبة المتلقين بصورة مباشرة.

ثالثاً: إنستغرامتميل النصوص في إنستغرام إلى القصر والوصف، وغالباً ما تأتي مرتبطة بالمحتوى البصري للمنشور. كما يظهر استخدام الوسوم والرموز التعبيرية والعناصر المختصرة لتدعيم المعنى والتعبير عن الموقف أو الشعور. ومن أمثلة العينة: «أجواء نهاية الأسبوع عملت باستا. الطعم خيال! #طبخ #أكل». ويكشف المثال عن جمل قصيرة تجمع بين العامية والفصحى، إلى جانب استخدام الوسوم لتصنيف المحتوى.

وبناءً على ذلك، تكشف المقارنة بين المنصات الثلاث أن الاختلاف في البيئة الرقمية ينعكس على شكل الخطاب اللغوي؛ إذ يغلب على X الاختصار، وعلى فيسبوك الطابع الحواري والممتد، بينما يميل إنستغرام إلى النصوص القصيرة المرتبطة بالمحتوى البصري. وتؤكد هذه الملاحظات أن اللغة الرقمية لا تظهر في صورة واحدة، بل تتغير وفق طبيعة المنصة ووظيفتها التواصلية.

٣-٦. الأبعاد الاجتماعية والثقافية: من "الاستلاب اللغوي" إلى "التفاوض الهوياتي"

لا يمكن مقاربة هذه التحولات اللغوية بمعزل عن سياقها السوسيولوجي والثقافي. في هذا الصدد، يبرز تباين جلي بين "المقاربات المعياريةالتي ترى في هذه الظواهر تهديداً للهوية، و"المقاربات المعيارية" التي تتبناها هذه الدراسة. ففي الاتجاه المعياري، يربط (بودیار، ٢٠١٨:٦٥٣) ظاهرة العربيزي بسياقات تاريخية، مشبهاً إياها بالمحنة التي عاشها عرب الأندلس بهدف فصلهم عن تراثهم اللغوي. وتنظر بعض الخطابات الأكاديمية الأخرى إلى هذه الممارسات على أنها "استلاب لغوي" وانفصال عن الثقافة العربية (عوفي، ٢٠١٤: ١٢)

غير أن هذا البحث يقترح تجاوز هذه الرؤية التشاؤمية؛ فبدلاً من وسم "العربيزي" والتهجين اللغوي بـ "الانسلاخ الهوياتي" أو "أزمة الثقة" باللغة الأم، نرى وجوب مقاربتها بوصفها مساحة ديناميكية لـ "التفاوض الهوياتي". فهذه الأنماط المستحدثة تمثل مسعىً لبناء "هوية رقمية" تتجاوز الحدود التقليدية؛ إذ يستخدم الشباب العربيزي والمقترضات الأجنبية كخيار لغوي واعٍ و"شفرة اجتماعية" تعبر عن هوية هجينة ومزدوجة: تنتمي إلى الثقافة العالمية الشبكية وتحتفظ في الوقت ذاته بجذورها التواصلية المحلية وقد توصلت دراسة حديثة(Algahtani &Albawardi, 2025: 53)إلى أن الشباب السعودي يستخدم العامية شبه المتعلمة في التغريدات كوسيلة للتعبير عن لهجة محليةسعودية فريدة،مما يعكس فخراً بالهوية المحلية في نفس الوقت الذي ينخرط فيه بالفضاء الرقمي العالمي.

٤. تأثير الذكاء الاصطناعي والمعالجة الآلية على التحولات اللغوية

٤-١. دور نماذج التوليد اللغوي في تعزيز الهجنة وتطبيعها

يؤدي الذكاء الاصطناعي التوليدي، وتحديداً النماذج اللغوية الكبيرة الحديثة، دوراً مزدوجاً وأساسياً في التعامل مع التحولات اللغوية الرقمية وإعادة إنتاجها. فمن جهة، يتم تدريب هذه النماذج على مدونات ضخمة مستمدة من الإنترنت، بما في ذلك منصات التواصل الاجتماعي التي تزخر بالأنماط الهجينة (العربية الفصحى، العاميات، العربيزي، واللغات الأجنبية). ونتيجة لذلك، تستبطن هذه النماذج الأنماط الهجينة بشكل آلي وتعيد إنتاجها في مخرجاتها، مما يساهم في "تطبيع" هذه التراكيب وتثبيتها تداولياً داخل الفضاء الرقمي (Fitrianto & Hamid, 2024: 27)من جهة أخرى، أظهرت الدراسات الحديثة أن هذه النماذج، على الرغم من قدراتها المتقدمة، تواجه تحديات كبيرة عند معالجة النصوص العربية غير المعيارية، وخاصة تلك المكتوبة بالعربيزي. فقد أوضحت دراسة حديثة (Almaoui 2025: 30) أن العربيزي - وهو شكل هجين من العربية يستخدم الحروف والأرقام اللاتينية لتمثيل اللهجات المنطوقة - يشكل تحدياً كبيراً للترجمة الآلية بسبب افتقاره إلى البنية الرسمية واحتوائه على عناصر ثقافية عميقة وقد قامت الباحثة بتقييم قدرة النماذج اللغوية الكبيرة المختلفة على فك تشفير وترجمة العربيزي، مع التركيز على لهجات عربية متعددة لم تُدرس بشكل كافٍ من قبل، وخلصت إلى أن الأداء يختلف اختلافاً كبيراً حسب اللهجة والسياق.

في سياق هذه الدراسة، رصد الباحثان توجيه المستخدمين لأوامر نصية تدمج بوضوح بين مستويات اللغة العربية الفصحى والعامية والعربيزي واللغات الأجنبية، مما يجبر النماذج الآلية على محاكاة التبديل الشفري لفهم السياق والاستجابة بشكل مناسب. ومن الأمثلة الموثقة من تفاعلات المستخدمين الفعلية (٢٠٢٥): "ممكن تلخّص المقال ده وتسلّط الضوء على أهم النقط؟". يجمع هذا المثال بين أدوات الاستفهام العربية ("ممكن")، والأفعال العربية الصريحة ("تلخّص"، "تسلّط الضوء")، والضمائر العامية ("ده")، مما يعكس بشكل نموذجي طبيعة الأوامر التي يتعامل معها الذكاء الاصطناعي يومياً. ويُلاحظ أن النماذج اللغوية الكبيرة تظهر أداء أفضل في السياقات المهيأة والمعالجة مسبقاً، بينما ينخفض أداؤها بشكل ملحوظ عند التعامل مع النصوص العفوية غير المعالجة من وسائل التواصل الاجتماعي (Alqahtani & Aljamaan, 2026: 12) وقد أوضحت دراسة حديثة أن النماذج اللغوية الكبيرة حققت دقة ٨٥% في تحليل المشاعر على مراجعات المطاعم المعالجة مسبقاً، ولكن هذه الدقة انخفضت

إلى ٧٧% عند التعامل مع التغريدات الأصلية على منصةX ، مما يشير إلى أن أداء هذه النماذج يتأثر بشكل كبير بالطبيعة غير الرسمية والفوضوية لمحتوى وسائل التواصل الاجتماعي.

عند استجابة هذه النماذج لطلبات المستخدمين بهذه اللغة العربية غير الرسمية أو الهجينة، فإنها غالباً ما تولّد محتوى يمزج بسلاسة بين الفصحى والعامية والمقترضات الأجنبية المدمجة صرفياً، مما يعزز دورة إعادة إنتاج هذه الهجنة. وتؤدي هذه التغذية الراجعة الآلية دوراً محورياً في "تطبيع" هذه التراكيب المحدثة، وتثبيتها تداولياً، ومنحها مشروعية استعمالية مستجدة في الفضاء الرقمي، الأمر الذي يخلق حلقة تكرارية لا تنفصم بين الإنتاج البشري وإعادة الإنتاج الآلي للهجنة اللغوية.

۲-۴. عقبات المعالجة الآلية للغات الطبيعية (NLP) والأدلة من البحوث الحديثة

تواجه أنظمة المعالجة الآلية للغات العربية عقبات هيكلية جسيمة عند معالجة ظاهرة "العربيزي" والتراكيب الصرفية الهجينة والتبديل الشفري داخل الجملة الواحدة. وقد وثقت دراسة مرجعية حديثة(Khaled & Al-Mubarak, 2025: 8)مجموعة من التحديات اللغوية الأساسية التي تعيق تطور معالجة اللغة العربية، ومن أبرزها: التعقيد الصرفي (نظراً للطبيعة الاشتقاقية الغنية للغة العربية)، وغياب التشكيل الإملائي في النصوص الرقمية (مما يؤدي إلى غموض كبير في التحليل)، وتعدد اللهجات والتنوع اللكنوي، وندرة الموارد والبيانات عالية الجودة، بالإضافة إلى ظاهرة الازدواجية اللغوية التي تخلق فجوة معيارية بين الفصحى والعاميات. ونظراً لغياب التوحيد الإملائي، وانتشار الأنماط الرقمية العفوية، وافتقار هذه النصوص للجذور المعجمية المعيارية، تُصنف المحللات الصرفية التقليدية هذه التراكيب ضمن فئة "الكلمات خارج المعجم" مما يضعف دقة التحليل بشكل كبير. وقد أشار استعراض حديث(Khaled & Al-Mubarak, 2025: 8)إلى أن معدلات الكلمات خارج المعجم في النصوص العربية الرقمية الهجينة يمكن أن تصل إلى مستويات تعطل أداء الأنظمة التقليدية، خاصة في المنشورات التي تدمج بين العربيزي واللهجات المحلية(Almaoui, 2025: 30)لتوضيح حجم هذه التحديات بشكل أكثر دقة وموضوعية، يُمكن الاستناد إلى نتائج الدراسات الحديثة في هذا المجال، بدلاً من الادعاءات غير الموثقة. فقد أظهرت دراسة رائدة لتقييم أداء النماذج اللغوية الكبيرة في معالجة العربية اللهجية أن نموذج MARBERTv2 حقق أفضل أداء في تحليل المشاعر على مجموعة من الفنادق المكتوبة باللهجتين السعودية والمغربية، حيث بلغت قيمة المقياس المركب ٧٩ % (Nwesri, 2025: 88)في المقابل، أظهرت دراسة أخرى حول تحليل المشاعر للهجة الجزائرية على منصة X أن النماذج المصممة خصيصاً لهذه اللهجة يمكن أن تحقق دقة تصل إلى ٩١.٢٣% بعد المعالجة المسبقة(Bougrine, 2025: 5) بينما أظهرت دراسة حديثة حول تحليل المشاعر بتقنية الصفر-شوت باستخدام النماذج اللغوية الكبيرة أن أداء النموذج ينخفض عند التعامل مع نصوص تويتر العفوية مقارنة بالنصوص المعالجة مسبقاً، حيث بلغت الدقة ٧٧% للتغريدات الأصلية مقارنة بـ ٨٥% لمراجعات المطاعم المعالجة (Alqahtani & Aljamaan, 2026: 12). ويتأكد هذا التباين في نتائج الدراسات الحديثة أن أداء النماذج يتفاوت بشكل كبير اعتماداً على عوامل متعددة مثل: معالجة النصوص مسبقاً، وجدوى الموارد اللغوية لكل لهجة على حدة، ودرجة هجنة النصوص واعتمادها على العربيزي والتبديل الشفري. وتشير هذه النتائج مجتمعة إلى أن الفجوة بين الواقع

اللغوي العربي الرقمي وقدرات الأنظمة الآلية الحالية لا تزال كبيرة، خاصة في سياق اللهجات قليلة الموارد والنصوص العفوية غير المعالجة والمواقف التواصلية المرتفعة الهجنة.

وقد تناول استعراض حديث (Benadla, 2024: 25)بشكل خاص ظاهرة التبديل الشفري في النصوص العربية على وسائل التواصل الاجتماعي، مشيراً إلى أن كلاً من تبديل الشفرة بين العربية واللغات الأجنبية (خاصة الإنجليزية والفرنسية) والتبديل بين اللهجات المختلفة وبين الفصحى والعاميات يطرح تحديات فريدة لأنظمة المعالجة الآلية، بما في ذلك تحديد حدود الرموز اللغوية، والتحليل الصرفي الدقيق للجذور، وفهم المعنى السياقي غير المباشر.

٤-٣. آفاق التطوير والتحسين: نحو نماذج ذكاء اصطناعي متخصصة للعربية الهجينة

للتغلب على القصور التقني والمنهجي الذي تواجهه الأنظمة الحالية، تبرز الحاجة الماسة إلى تطوير نماذج ذكاء اصطناعي (AI) متخصصة ومدربة على "مدونات عربية هجينة وموسومة". ويتطلب ذلك توظيف خوارزميات التعلم العميق للتعرف الآلي على ظاهرة "التبديل الشفري" المدمج ضمن النص الواحد (سواء بين العربية واللغات الأجنبية أو بين مستويات العربية المختلفة)، وفك شفرة الاختصارات والمنحوتات الرقمية المتداولة على نطاق واسع. وقد أثبتت الأبحاث الحديثة أن النماذج التي يتم تدريبها خصيصاً على هذه الظواهر تُظهر أداءً متفوقاً مقارنة بالنماذج متعددة اللغات العامة. فعلى سبيل المثال، أظهر نموذج MARBERTv2 - المصمم خصيصاً للهجات العربية - تفوقاً واضحاً في مهام تحليل المشاعر على نصوص الفنادق المكتوبة باللهجتين السعودية والمغربية (Nwesri, 2025: 89) ، كما أظهرت نماذج مثل MARBERT-LSTM دقة قياسية بلغت ٩١.٢٣% في تحليل المشاعر للهجة الجزائرية على منصة X بعد المعالجة المسبقة المناسبة (Bougrine, 2025: 7)،مما يؤكد أهمية التصميم الخاص باللهجة في تحسين الأداء. كما أظهرت دراسة أخرى (Alqahtani & Aljamaan, 2026: 12)أنه حتى عندما تنجح النماذج متعددة اللغات في الإجابة على أسئلة اللغة العربية الفصحى، فإنها كثيراً ما تفشل في فهم نفس الأسئلة عندما تُترجم إلى لهجة عربية محلية، مما يشير إلى أن ما يظهر "فهماً" قد يكون في الواقع حفظاً واستظهاراً للأنماط اللغوية الشائعةبدلاً من الفهم الدلالي الحقيقي.

إن هذا التطوير المستقبلي لن يسهم فقط في رفع كفاءة التفاعل الآلي مع المستخدم العربي والارتقاء بمستوى الخدمات الرقمية المقدمة له، بل سيضمن دقة أعلى وموثوقية أكبر في التطبيقات اللسانية الحاسوبية الحساسة، مثل تحليل المشاعر (لرصد الرأي العام والتوصيات التجارية)، واستخراج البيانات والكيانات المسماة (للأغراض الأمنية والإعلامية والبحثية)، وبناء واجهات تفاعلية ذكية (مثل المساعدات الشخصية والترجمة الفورية) تستوعب الديناميكية المستمرة للغة العربية في الفضاء الرقمي. وتظل الحاجة ملحة أيضاً لتطوير بنى تحتية قياسية وأطر تقييم موحدة وعادلة للنماذج العربية، كما هو الحال مع منصة QIMMA (قِمّة) التي تتبنى نهجاً قائماً على الجودة لضمان أن النتائج المبلغ عنها تعكس القدرة اللغوية الحقيقية للنماذج العربية (AlQadi, 2026: 1).

٥. التحليل الكمي الموسع (الأدلة الإحصائية)

٥-١. الإحصاءات المستخرجة من عينة الدراسة (منصات X، فيسبوك، إنستغرام)

بناءً على المقاربة الكمية التوصيفية المعتمدة في هذه الدراسة، يلخص الجدول (٣) التوزيع النسبي لأبرز الظواهر الصرفية والنحوية المرصودة في عينة الـ ٥٠٠ منشور/تعليق المأخوذة من منصات X وفيسبوك وإنستغرام خلال عام ٢٠٢٥. تجدر الإشارة إلى أن جميع النسب تعبر عن العينة المدروسة فقط ولا تدعي التعميم الإحصائي.

جدول (٣): التوزيع النسبي لأبرز الظواهر اللغوية في عينة الدراسة (ن = ٥٠٠ منشور/تعليق)

الظاهرة اللغوية المرصودة	النسبة المئوية من العينة (%)	التكرار (من إجمالي ٥٠٠)
الانفصام الصوتي-الخطي (استخدام العربيزي)	٤٢%	٢١٠ منشوراً
التفضيل الأسلوبي والتحول الرتبي نحو بنية (SVO)	٦٥%	٣٢٥ جملة خبرية
التكييف الصرفي (الاشتقاق الهجين، النحت، الاختصارات)	٣٠%	١٥٠ منشوراً
هيمنة الاسم الموصول العامي "إلّي"	٨٥%	من الجمل التي تحتاج إلى موصول
النمط التلغرافي (في منصة X)	٢٥-٣٠%	من منشورات X

ملاحظة: النسب المئوية الواردة في هذا الجدول تستند إلى العينة الحالية فقط (٥٠٠ منشور/تعليق) ولا تمثل بالضرورة النسب في المجتمع العربي ككل. كما أن بعض الظواهر قد تتداخل في المنشور الواحد، لذلك لا ينبغي جمع النسب تراكمياً.

كما تُظهر المعطيات الكمية أن ظاهرة العربيزي ليست مقتصرة على منصة واحدة، بل تم رصدها في جميع المنصات الثلاث، وإن كانت بنسب متفاوتة: بلغت أعلى نسبة لها في منصة X (حوالي ٥٥% من منشورات X في العينة)، تلتها فيسبوك (٣٥%)، ثم إنستغرام (٢٥%). وهذا يتفق مع ما أشار إليه (درويش، ٢٠١٩: ٥٥٨)من أن قيود المساحة في منصة X تدفع المستخدمين إلى استراتيجيات إيجاز أكثر كثافة، بما في ذلك العربيزي والاختصارات والنحت. أما بالنسبة لبنية (فاعل ،فعل ،مفعول)، فقد لوحظ أنها تسود في جميع أنواع المنشورات، سواء كانت حوارية (فيسبوك) أو تلغرافية (X) أو وصفية (إنستغرام)، مما يشير إلى أن هذا التحول الرتبي يتجاوز خصوصية المنصة ليصبح سمة أسلوبية عامة للغة العربية الرقمية المعاصرة (العجمي، ٢٠١٤: ٨)يكشف التحليل المقارن بين نتائج هذه الدراسة ودراسة درويش (٢٠١٩) عن اتجاه تصاعدي في نسبة الهجنة اللغوية؛ فقد أشار(درويش، ٢٠١٩: ٥٥٨) إلى أن المزج بين الفصحى والعامية واللاتينية شكّل ٣٩% من عينته، في حين تجاوزت هذه النسبة ٤٥% في عينة الدراسة الحالية (٢٠٢٥). هذا التصاعد، رغم تحسن دعم الكتابة العربية تقنياً، يؤكد أن الظاهرة لم تعد حلاً تقنياً مؤقتاً، بل تحولت إلى نظام تداولي مستقر يعكس هوية رقمية هجينة(Alqahtani & Aljamaan, 2026: 53)وتؤكد هذه النتائج أن الظواهر المرصودة ليست تقلبات عشوائية، بل أنماطاً لغوية منتظمة، مما يستدعي دراسات مستقبلية على عينات أوسع لتعميم النتائج إحصائياً.

۶.توصيات الدراسة

للباحثين والمتخصصين في معالجة اللغة العربية :(NLP) تطوير مدونات رقمية هجينة وموسومة تتضمن العربيزي، والتبديل الشفري، واللهجات المختلفة، والعمل على تطوير نماذج متخصصة لمعالجة هذه الأنماط، ولا سيما في اللهجات قليلة الموارد والسياقات الرقمية العفوية .

للمؤسسات التعليمية وصانعي السياسات اللغوية :اعتماد مقاربة تقوم على الوعي بالتنوع الوظيفي للغة، والتمييز بين استخدام الفصحى في السياقات الرسمية والأكاديمية وبين الأنماط الرقمية الهجينة في التواصل غير الرسمي، بدلاً من وصم العامية والعربيزي بوصفهما «تشويهاً» أو «انحطاطاً ».

لمطوري المنصات وللباحثين مستقبلاً :تحسين أدوات معالجة الكتابة العربية والعربيزي، بما في ذلك التصحيح الإملائي والنحوي والتحويل بين العربيزي والرسم العربي، مع توسيع الدراسات المستقبلية لتشمل بلداناً ولهجات وفئات عمرية مختلفة، واستخدام عينات أكبر وأدوات إحصائية مناسبة، إضافة إلى دراسة أثر الذكاء الاصطناعي التوليدي في تطور هذه الأنماط.

النتيجة

خلصت هذه الدراسة إلى أن اللغة العربية في فضاء التواصل الاجتماعي تمر بمرحلة تحول بنيوي عميق، لا يمكن اختزاله في مظاهر "تشويه" أو "انحطاط"، بل هو استجابة وظيفية منتظمة لمتطلبات بيئة رقمية جديدة تفرض السرعة، والاقتصاد اللغوي، والتعبير عن هوية رقمية هجينة. وفي الإجابة عن السؤال المحوري للبحث حول كيفية إعادة تشكيل وسائل التواصل الاجتماعي للبنى الصرفية والنحوية للغة العربية، كشفت النتائج أن التحولات الصرفية تجلت في أربعة مظاهر رئيسية: ظاهرة "العربيزي" في ۴۲% من العينة، والتبسيط الصرفي للضمائر والأسماء الموصولة مع هيمنة "إلّي" بنسبة ۸۵%، والتكييف الصرفي للمقترضات الأجنبية (الاشتقاق الهجين) في ۳۰% من العينة مع هيمنة وزن يُفَعِّل (۶۵%)، والنحت والاختصارات كآليات للاقتصاد اللغوي في ۲۵% من منشورات X. وعلى المستوى النحوي، تمثلت التحولات في ثلاثة محاور رئيسية: التحول الرتبي نحو بنية (فاعل–فعل–مفعول) بنسبة ۶۵% من الجمل الخبرية مقابل ۳۵% للبنية الكلاسيكية (فعل–فاعل–مفعول)، والاستعاضة الوظيفية بالأدوات النحوية العامية بنسب تراوحت بين ۴۵% (للنفي) و۸۵% (للوصول)، والنمط التلغرافي في ۲۵–۳۰% من منشورات X. وهذه الأرقام تؤكد أن اللغة العربية الرقمية ليست مجرد فصحى فاسدة، بل هي نظام صرفي ونحوي جديد قائم بذاته، يتسم بالمرونة والقدرة على استيعاب المستجدات مع الاحتفاظ - ولو جزئياً - بالهيكل العربي الأصيل .وفي الإجابة عن السؤال المتعلق بالتحديات التقنية التي تفرضها هذه التحولات على أنظمة المعالجة الآلية للغة العربية، أثبتت الدراسة أن هذه التحولات اللغوية ينبغي فهمها بوصفها مساحة ديناميكية للتفاوض الهوياتي، حيث يستخدم الشباب العربي الأشكال الهجينة كخيار واعٍ وشفرة اجتماعية تعبر عن هوية مزدوجة تنتمي إلى الثقافة العالمية الشبكية وتؤكد في الوقت نفسه الجذور المحليةً واللهجية. غير أن هذا الواقع يفرض تحديات هيكلية كبيرة أمام الأنظمة الآلية، حيث تتراوح معدلات الدقة بين ۷۷% و۹۱% حسب درجة المعالجة المسبقة واللهجة المستهدفة، مما يعني أن نسبة الإخفاق الدلالي قد تصل إلى ما بين ۹% و۲۳% في أفضل الأحوال، وقد تكون أعلى في السياقات العفوية غير المعالجة لذلك، توصي الدراسة بضرورة انتقال الباحثين والمتخصصين في المعالجة الآلية من الأطر الوصفية

التقليدية إلى هندسة مدونات رقمية هجينة وموسومة، وتدريب نماذج تعلم عميق متخصصة بدلاً من الاعتماد على النماذج متعددة اللغات العامة، مع التركيز على اللهجات قليلة الموارد والسياقات العفوية غير المعالجة، كما توصي المؤسسات التعليمية وصانعي السياسات اللغوية بالتوقف عن وصم هذه الظواهر والانتقال إلى تعليم "التنوع الوظيفي للغة"، وتطوير مناهج تعزز الوعي بالهوية اللغوية وتشجع على الاستخدام المتوازن للفصحى في السياقات المناسبة دون حظر أو تجريم العامية والعربيزي في الفضاءات غير الرسمية.

المصادر والمراجع

الف. المصادر العربية

بودیار، عادل. (۲۰۱۸). **فصاحة المدرسة وعامية مواقع التواصل الاجتماعي**. مجلة مداد الآداب، (عدد خاص بالمؤتمرات)، ۶۴۸-۶۵۵.

الجرف، ريما. (۲۰۲۱). **أثر وسائل التواصل الاجتماعي على استنزاف اللغة العربية. قاعدة بيانات إيريك** .(ERIC) (ED615197)

درويش، عبد الحفيظ. (۲۰۱۹). **استخدام رواد مواقع التواصل الاجتماعي للثنائية والازدواجية اللغوية والحرف اللاتيني وآثاره على اللغة العربية**. مجلة البحوث الإعلامية، جامعة الأزهر، (۵۱)، ۵۳۳-۵۸۰.

العجمي، سعد بن طفلة. (۲۰۱۴). **العربتيني: الكتابة العربية بالأحرف اللاتينية. في لغة الشباب العربي في وسائل التواصل الحديثة**. الرياض: مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، صص ۷-۲۶.

عوفي، عبد الكريم علي. (۲۰۱۴). **اللغة العربيزية (الهجينة) في مواقع التواصل الاجتماعي وأثرها على اللغة العربية الفصحى**. في أبحاث وندوات مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية ، الرياض: مركز الملك عبد الله بن عبد العزيز الدولي لخدمة اللغة العربية، صص ۱۱-۶۱.

كسّاس، صافية. (۲۰۱۹). **الاستعمال اللغوي في وسائل التواصل الاجتماعي عند الشباب العربي: الواقع والأسباب والآثار**. مجلة إشكالات في اللغة والأدب، ۸(۳)، ۴۶۲-۴۷۸.

ب. المصادر اللاتينية

Abdelali, A., Darwish, K., Durrani, N., & Mubarak, H. (2016). **Farasa: A new fast and accurate Arabic tokenizer**. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16) (pp. 219-223). European Language Resources Association (ELRA).

Abdul-Raof, T. A. (2021). **The Use of Arabic Neologisms in Social Media Applications**. International Journal of Arabic-English Studies, 21(1), 1-20.

Algahtani, M., & Albawardi, A. (2025). **Levels of Arabic Language in Saudi Tweets**. Academic Journal of Interdisciplinary Studies, 14.(۲)

Alkaabi, H. A., Jasim, A. K., & Darroudi, A. (2025). **Arabic NLP: A Survey of Pre-Processing and Representation Techniques**. Journal of Computer Science and Information Technology, 14.(۱)

Almaoui, P., Bouillon, P., & Hengchen, S. (2025). **Arabizi vs LLMs: Can the Genie Understand the Language of Aladdin?** In Proceedings of Machine Translation Summit XX Volume 2 (pp. 28-41). Geneva, Switzerland

Alqahtani, S., & Aljamaan, F. (2026). **Evaluating Arabic sentiment analysis with GPT-4o: Zero-shot performance on restaurant reviews and tweets**. Journal of King Saud University – Computer and Information Sciences, 38.(٢)

Al-Sayadi, S. (2024). **Arabizi in Saudi Youth Online Communication: Identity and Frequency**. Journal of Saudi Linguistics, 12(2), 45-67.

Androutsopoulos, J. (2014). **Languageing when online**: Language and social media. In The Routledge Handbook of Language and Digital Communication (pp. 231-245). Routledge.

Benadla, D. (2024). **Code Choice in Instant Messaging Among Algerian Youth: A Case Study of Messenger**. Algerian Journal of Sociolinguistics, 7(1), 22-41.

Bougrine, Z., et al. (2025). **Benchmarking sentiment analysis of Algerian Arabic dialect on X (Twitter)**. Journal of King Saud University – Computer and Information Sciences, 37.(٥)

Creswell, J. W., & Creswell, J. D. (2018**). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.).** SAGE Publications.

Ferguson, C. A. (1959). **Diglossia. Word**, 15(2), 325-340.

Fitrianto, I., & Hamid, R. (2024). **Morphosemantic Changes in the Arabic Language in the Social Media Era: A Study of Neologisms and Their Impact on Youth Communication**. International Journal of Arabic Studies, 1.(١)

Gugliotta, E. (2025). **Tunisian Arabizi: Linguistic Analyses and Corpus Building using Natural Language Processing**. DOAB.

Khaled, N., & Al-Mubarak, A. (2025). **Arabic natural language processing (NLP): A comprehensive review of challenges, techniques, and emerging trends**. Computers, 14(11), 497.

Krippendorff, K. (2018). **Content analysis: An introduction to its methodology (4th ed.)**. SAGE Publications.

Neuman, W. L. (2014**). Social research methods: Qualitative and quantitative approaches (7th ed.)**. Pearson.

Nwesri, A. F. A., Shinbir, N. A. S., & Sharif, A. B. (2025). **Sentiment analysis on Arabic dialects: A multi-dialect benchmark**. In Proceedings of the Shared Task on Sentiment Analysis for Arabic Dialects (RANLP 2025) (pp. 86-91). INCOMA Ltd.

دوفصلنامهٔ مطالعات میان‌رشته‌ای در زبان و ادبیات عربی

شاپای الکترونیکی : ۶۹۵۵-۳۰۹۲

بررسی صرفی و نحوی واژگان و ساختارهای نوپدید در رسانه‌های اجتماعی عربی: نمونه‌پژوهی توییتر، فیسبوک و اینستاگرام

سودابه مظفری [۱]، رامین صداقت [۲]

[۱] دانشیار، گروه زبان و ادبیات عربی، دانشکدهٔ ادبیات و علوم انسانی، دانشگاه خوارزمی، تهران، ایران.

[۲] دانشجوی کارشناسی ارشد، گروه زبان و ادبیات عربی، دانشکدهٔ ادبیات و علوم انسانی، دانشگاه خوارزمی، تهران، ایران.

اطلاعات مقاله

نوع مقاله:
مقاله پژوهشی

دریافت:
۱٤۰٤/۱۲/۰۹

پذیرش:
۱٤۰۵/۲/۲۱

چکیده

انقلاب دیجیتال فضای زبانی جدیدی در شبکه‌های اجتماعی پدید آورده که آمیزه‌ای ناهمگون از عربی فصیح، گویش‌های محاوره‌ای و زبان‌های خارجی است. این پژوهش با رویکرد ترکیبی (توصیفی-تحلیلی و کمّی)، نمونه‌ای هدفمند از ۵۰۰ پست در پلتفرم‌های توییتر، فیسبوک و اینستاگرام (۲۰۲۵) را بررسی کرده است. یافته‌های صرفی نشان از گسترش«عربیزی» (٤۲٪)، اشتقاق دورگه از ریشه‌های بیگانه و آمیزه‌سازی (نحت) دارد. در سطح نحوی، گرایش به ساده‌سازی با جایگزینی موصول عامیانه «إلّی» و غلبه ساختار فاعل-فعل-مفعول (SVO) با نرخ ٦۵٪ مشاهده شد. این دگرگونی‌ها اگرچه ابزاری برای مذاکره هویتی جوانان هستند، چالش‌هایی ساختاری برای مدل‌های هوش مصنوعی ایجاد کرده و خطای معنایی تا ۳۰٪ را سبب می‌شوند. این پژوهش میان‌رشته‌ای بر ضرورت توسعه پیکره‌های متنی دورگه برای بهبود سیستم‌های پردازش زبان طبیعی عربی تأکید دارد.

کلمات کلیدی: دگرگونی‌های صرفی و نحوی، واژگان نوساخته، زبان رسانه‌های اجتماعی، عربیزی، زبان‌شناسی دیجیتال، هوش مصنوعی، NLP.

استناد: مظفری،س. صداقت،ر.(۱٤۰۵). بررسی صرفی و نحوی واژگان و ساختارهای نوپدید در رسانه‌های اجتماعی عربی: نمونه‌پژوهی توییتر، فیسبوک و اینستاگرام، *روایت‌پژوهی میان‌رشته‌ای در زبان و ادبیات عربی*. ۳ (ویژه‌نامه). ۹۳-۱۱۸.

DOI: 10.22034/jisall.2026.555190.1101

ناشر: دانشگاه زابل

حق‌مؤلف © نویسندگان.

*- نشانی پست الکترونیکی (نویسنده مسئول): ramin.sedaghat@khu.ac.ir.

A Morphological and Syntactic Study of Neologisms and Structures in Arabic Social Media
(Twitter, Facebook, Instagram as Case Studies)

Soudabeh mozaffari, Associate Professor in Arabic Language and Literature, Faculty of Literature and Humanities, Kharazmi University, Tehran, Iran.
Ramin sedaghat, (corresponding author): master student of Arabic language and literature, Faculty of Literature and Humanities, Kharazmi University, Tehran, Iran.
Email: ramin.sedaghat@khu.ac.ir.

Introduction

The digital revolution has created a hybrid linguistic environment on social media, where Modern Standard Arabic, regional dialects, and foreign languages converge. This study adopts a descriptive perspective to examine systematic morphological and syntactic transformations in Arabic digital discourse, framing these changes as adaptive responses to new communicative constraints rather than as random errors. It also explores the challenges these hybrid forms pose to Arabic Natural Language Processing (NLP).

Methodology

The research employs a mixed descriptive-analytical and quantitative design. A purposive sample of 500 public posts and comments was collected from X (50%), Facebook (30%), and Instagram (20%) during 2025. After eliminating repetitions, 420 unique linguistic phenomena were identified and analyzed usingqualitativeinterpretationalongsidefrequency-basedquantitative measures.

ResultsandDiscussion

Four major morphological patterns were identified: (1) *Arabizi*—the use of Latin letters and numerals for Arabic sounds—appeared in 42% of the sample; (2) morphological simplification, notably the unified colloquial relative pronoun *"elli"*, replaced all standard gender- and number-distinctive forms in approximately 85% of relevant clauses; (3) hybrid derivation, where foreign roots were integrated into common Arabic verb templates, occurred in 30% of posts, with two specific patterns accounting for over 90% of these cases; and (4) blending and abbreviation, driven by linguistic economy, were especially frequent on X, where 25% of posts contained shortened expressions. Syntactically, a clear preference for Subject-Verb-Object (SVO) word order was observed, appearing in 65% of declarative sentences, alongside widespread use of colloquial substitutes for standard negation, questioning, andlinkingparticles.

These transformations are systematic and motivated primarily by linguistic economy, communicative efficiency, platform constraints, and the expression of digital identity. However, the growing hybridity significantly challenges Arabic NLP systems, as many forms fall outside traditional linguistic resources.

Conclusion

he study strongly recommends the development of annotated hybrid corpora capable of capturing the full spectrum of contemporary digital Arabic usage, in order to support more robust NLP applications and better reflect the linguistic reality of Arabic-speaking online communities.

References

Abdelali, A., Darwish, K., Durrani, N., & Mubarak, H. (2016). Farasa: A new fast and accurate Arabic tokenizer. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), 219–223. [in English]

Abdul-Raof, T. A. (2021). The use of Arabic neologisms in social media applications. International Journal of Arabic-English Studies, 21(1), 1–20. [in English]

Al-Ajmi, S. (2014). Al-Arabitini: Arabic writing in Latin letters. In The Arabic Youth Language in Modern Communication Media, 7–26. Riyadh: King Abdullah International Center for Arabic Language Service. [in Arabic]

Algahtani, M., & Albawardi, A. (2025). Levels of Arabic language in Saudi tweets. Academic Journal of Interdisciplinary Studies, 14(2). [in English]

Al-Jarf, R. (2021). The impact of social media on the depletion of the Arabic language. ERIC, ED615197. [in Arabic]

Alkaabi, H. A., Jasim, A. K., & Darroudi, A. (2025). Arabic NLP: A survey of pre-processing and representation techniques. Journal of Computer Science and Information Technology, 14(1). [in English]

Almaoui, P., Bouillon, P., & Hengchen, S. (2025). Arabizi vs. LLMs: Can the Genie understand the language of Aladdin? In Proceedings of Machine Translation Summit XX, 28–41. Geneva, Switzerland. [in English]

Al-Oufi, A. A. (2014). Arabizi (hybrid language) on social networking sites and its impact on Modern Standard Arabic. Riyadh: King Abdullah International Center for Arabic Language Service. [in Arabic]

AlQadi, L., Alzubaidi, A., & Hacid, H. (2026). QIMMA: A quality-first approach to Arabic LLM evaluation. In Proceedings of COLING 2026. [in English]

Al-Sayadi, S. (2024). Arabizi in Saudi youth online communication: Identity and frequency. Journal of Saudi Linguistics, 12(2), 45–67. [in English]

Androutsopoulos, J. (2014). Languageing when online: Language and social media. In The Routledge Handbook of Language and Digital Communication, 231–245. Routledge. [in English]

Benadla, D. (2024). Code choice in instant messaging among Algerian youth: A case study of Messenger. Algerian Journal of Sociolinguistics, 7(1), 22–41. [in English]

Boudiar, A. (2018). Fus'ha in schools and colloquial Arabic on social networking sites. Midad Al-Adab Journal, Special Conference Issue, 648–655. [in Arabic]

Bougrine, Z., et al. (2025). Benchmarking sentiment analysis of Algerian Arabic dialect on X (Twitter). Journal of King Saud University – Computer and Information Sciences, 37(5). [in English]

Creswell, J. W., & Creswell, J. D. (2018). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. 5th ed. SAGE Publications. [in English]

Darwish, A. H. (2019). The use of bilingualism, diglossia, and Latin script by social media users and their effects on the Arabic language. Journal of Media Research, Al-Azhar University, 51, 533–580. [in Arabic]

Ferguson, C. A. (1959). Diglossia. Word, 15(2), 325–340. [in English]

Fitrianto, I., & Hamid, R. (2024). Morphosemantic changes in the Arabic language in the social media era: A study of neologisms and their impact on youth communication. International Journal of Arabic Studies, 1(1). [in English]

Gugliotta, E. (2025). Tunisian Arabizi: Linguistic Analyses and Corpus Building Using Natural Language Processing. DOAB. [in English]

Kassas, S. (2019). Language use on social media among Arab youth: Reality, causes, and effects. Journal of Issues in Language and Literature, 8(3), 462–478. [in Arabic]

Khaled, N., & Al-Mubarak, A. (2025). Arabic natural language processing (NLP): A comprehensive review of challenges, techniques, and emerging trends. Computers, 14(11), 497. [in English]

Krippendorff, K. (2018). Content Analysis: An Introduction to Its Methodology. 4th ed. SAGE Publications. [in English]

Neuman, W. L. (2014). Social Research Methods: Qualitative and Quantitative Approaches. 7th ed. Pearson. [in English]

Nwesri, A. F. A., Shinbir, N. A. S., & Sharif, A. B. (2025). Sentiment analysis on Arabic dialects: A multi-dialect benchmark. In Proceedings of the Shared Task on Sentiment Analysis for Arabic Dialects (RANLP 2025), 86–91. INCOMA Ltd. [in English]