حوسبة العامية إلى الفصيحة ضمن مشروع تطوير تطبيقات لتصحيح اللغة العامية الأردنية إلى لغة عربية سليمة بتقنيات تعلم الآلة*

دلال إبراهيم الزهيري[1]، حسين فتحي عدوان[2]، غيث علي عبندة[3]، يوسف حسين حمدان[4]

[1،2،4] قسم اللغة العربية وآدابها

[3] قسم هندسة الحاسوب

الجامعة الأردنية، عمان، الأردن

[1] dalalalzuheiri@gmail.com, [2] husseinadwan71@gmail.com, [3] abandah@ju.edu.jo

[4] y.hamdan@ju.edu.jo

الخلاصة: يعالج هذا البحث جانبًا من القضايا المتعلقة بتحويل اللهجة العامِّيَّة المنطوقة أو المكتوبة إلى اللغة العربية الفصحى، وذلك عبر تقنيات الذكاء الاصطناعي وتعلُّم الآلة ضمن مشروع "تطوير تطبيقات لتصحيح اللغة العامية الأردنية إلى لغة عربية سليمة بتقنيات تعلم الآلة". وقد توقَّف المشروع في سبيل تحقيق ذلك على بُعدين متكاملين؛ البعد اللساني الحاسوبي: وفيه نقدم وصفًا تكنولوجيًّا لآليات المشروع ودور الحاسوب في معالجة اللغة بواسطة تقنيات تعلم الآلة العميق، والبعد اللغوي: وفيه نقدم نماذجَ على التَّداخلات اللُّغويَّة بين اللهجة العامِّيَّة واللغة الفصحى السَّليمة مع ملاحظة التَّحويرات التي نجم عنها هذا التداخل على مستوى الصوت والصَّرف والإملاء والتَّركيب والبلاغة والدَّلالة. علاوة على ملاحظات ختاميَّة عامَّة كشفت عنها طبيعة النَّصوص التي اشتغل عليها المشروع. ويميِّز هذا المشروع اشتغاله على معالجة النُّصوص الحيَّة الناتجة عن الواقع اللغوي في المجتمع متَّخذًا من اللَّهجة الأردنية الدَّارجة نموذجًا كلِّيًا مقابلًا للغة الفصحى بهدف تطوير تطبيقات حاسوبية تسهم في المحافظة على نشر محتوى لغوي فصيح وسليم سواء على مستوى التواصل اليومي أو في المحاورات والمحادثات المكتوبة عبر الوسائل التقنية التي يستعملها المجتمع بكل مستوياته. ولعلَّ النَّجاح الذي ستقيمه هذه التطبيقات في اللهجة الأردنية يوسِّع رقعة العمل لمشروعات مُشابِهة على مستوى الأقطار العربيَّة كافَّة، مما يسهل من سبل الاتصال الثقافي.

الكلمات الجوهرية: اللغة العربية، حوسبة، ترجمة، لهجة عامية أردنية، تعلم الآلة.

1. المقدمة

تعدُّ قضيَّة معالجة اللهجة العامِّيَّة –بصفتها المستوى اللغويّ الأول الذي تحاكيه الملكة اللغوية عند الإنسان في طفولته– وتحويلها إلى الفصحى بواسطة تقنيات الآلة والذكاء الاصطناعي حاجةً ملحّة؛ لأنَّ المتكلم كثيرًا ما يحتاج إلى التعبير

*مشروع مدعوم من صندوق دعم البحث العلمي والابتكار الأردني

[3] غيث عبندة في إجازة تفرغ علمي يقضيها كأستاذ زائر في الجامعة الأمريكية في الشارقة

بالفصيحة في المنصات الرَّقمية العالميَّة المختلفة، ولكنه لا يجد سبيلًا إلى ذلك لعوامل عديدة؛ فإمَّا أن يقع في كثير من الأخطاء في أثناء تعبيره، أو يخلط بين العامِّيَّة والفصحى، أو يلجأ إلى الكتابة بالعامية مباشرة. ومن شأن هذه العوامل أن تُضعف من عملية التعبير والفهم، علاوة على الضعف الذي ينسحب على المحتوى العربي في العالم. من هنا تأتي ميزة العمل الذي يقدمه هذا المشروع بحكم أنه يحلل نصوصًا من اللهجة العامية الأردنية ويردها إلى مقابلها الفصيح الذي تتقنه الآلة بنسبة ما، مع الحفاظ على المقاربة في البنية بين الفصيح والعامِّيّ، وهذا من شأنه أن يسهم في رفع الكفاءة التواصلية لمتكلمي العربية ومتعلميها؛ إذ يتيح لهم التعرف إلى النموذج الفصيح الذي تطرحه العربية من دون الوقوع في الأخطاء الإملائية والتركيبية والدلالية، إضافة إلى أنه يساعد الوافدين إلى تعلم العربية على فهم المستوى العامي الذي يتفرع عن الفصيح فيوضح لهم بعض الأبعاد الفكرية التي تكتنف المحكيات العامية ولا سيما في الأمثال والأقوال الشعبية؛ إذ ستتمكن الآلة من تقديم ترجمة سليمة تتماشى مع الفكرة الصحيحة التي تحكيها العامية، وهو ما سيسهِّل من تداول العربية في أغراض وسياقات مختلفة. ويشتمل هذا العمل على التَّأطير لبُعدين؛ بعد لسانيّ حاسوبيّ يبيِّنُ آلية العمل في المشروع، وبعد لغوي يُبيِّن مستويات التَّداخل بين العامِّيَّة والفصحى.

2. البُعدُ اللِّسانيّ الحَاسوبيّ

يهيمن الذكاء الاصطناعي على معظم مجالات حياتنا في الوقت الحالي؛ فلم يقتصر على الصناعات وإنما تعدَّاها إلى اللُّغة عبر الاهتمام بحوسبة اللغة، فغدا من صُلب اهتماماته "إنشاء برامج تستوعب اللغة الطبيعية فهمًا، وتطلقها مخرجات تقنية تُجارِي قدرة الإنسان على التعامل معها، وهو ما بات يُعرف باللسانيات الحاسوبية أو المعالجة الآلية للغة، أو هندسة اللغة، وجميعها مسمَّيات لمبحث لسانيّ جديد يجمع بين اللغة والحاسوب لمعالجة اللغة الإنسانية مُعالجة آلية [1].

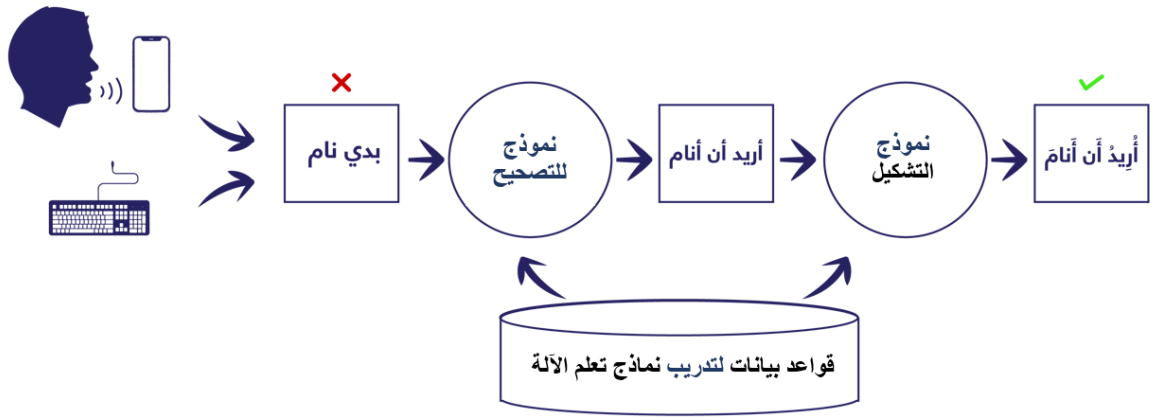
وهذا المشروع ينطلق معتمدًا على تقنيات الحوسبة، ويتميز في أنه يتجاوز مستوى اللغة الفصيحة الذي تعالجه معظم التطبيقات الحاسوبية في الوقت الحالي إلى معالجة مستوى العامِّيَّة المحكيَّة في إحدى اللهجات العربية؛ اللهجة الأردنيَّة. وفي سبيل ذلك ارتأى مبدأ المحاكاة إذ يستدعي نصوصًا لغوية حية من العامية المتداولة في المجتمع الأردني من مواقع التواصل الاجتماعي بمختلف طبقاته الاجتماعية، فضلًا عن الأفلام والمسلسلات الأردنية العامية.

ويرمي في نهاية المطاف إلى مساعدة المتكلم للوصول إلى مستوى جيد من اللغة الفصيحة إذا ما احتاج إلى ذلك في سياق التعلم أو الكتابة الصحيحة نحوًا وإملاءً من خلال تطوير تطبيقات حاسوبية وتوفيرها لمتحدثي اللغة العربية في الأردن تمكّنهم من تصحيح النصوص العربية التي يُدخلونها عبر حواسيبهم وهواتفهم المحمولة وتحويلها من نصوص باللهجة الأردنية العامية أو من نصوص فيها أخطاء لغوية مختلفة إلى نصوص عربية سليمة ومُشكّلة تشكيلًا كاملًا. ومما لا ريب فيه أنَّ توفير هذه التطبيقات سيكون له أثرٌ كبيرٌ في تسهيل الاتصال بين مستخدمي اللغة العربية في الأردن وإزالة كثير من أخطاء الاتصال واللُبس والارتقاء بلغة الاتصال وتحسين المحتوى العربي. كذلك سيكون لاستخدام هذه التطبيقات أثر إيجابي في تحسين المستوى اللغوي عند المستخدِمين وذلك عندما يرى المستخدِم كيف تُصَحِّحُ هذه التطبيقات أخطاءه اللغوية أولًا بأول. علاوة على ذلك، سيقدِّمُ المشروع إسهامات بحثية علمية في توفير قواعد بيانات لازمة لأبحاث اللغة العربية، وفي تطوير خوارزميات لمعالجة اللغات الطبيعية واستخدام تقنيات تعلم الآلة لتصحيح الأخطاء الإملائية واللغوية المختلفة. ونأمل في أن تسهم النواتج العلمية لهذا المشروع في تقدم الدعم الحاسوبي للغة العربية في الأردن وفي العالم العربي عمومًا.

ونسعى إلى نشر بحث متكامل عن هذا المشروع يتضمَّن أهداف المشروع وآلياته، وقواعد البيانات التي جَمَعَها واستعملها في عمليات تعلم الآلة، ونماذج وتجارب تعلم الآلة المستعملة، وتحليلات للنتائج، ووصفًا لتطبيقات الإنترنت والهواتف المحمولة المطوَّرة والتي ستُوفَّر للمستخدمين في الأردن.

ولأنَّ "تمثيل المعارف البشرية في الحاسبات الإلكترونية شبيه بتمثيل المعارف اللغوية في الدماغ البشري، وخاصة التمثيل الدّلالي، والبحث الحالي لعلم اللسانيات الحاسوبية هو تطبيق برامج آلية تمكنها من فهم اللغات البشرية" [1]، فقد ارتأى المشروع تعليم الآلة بالمحافظة على النموذج اللغوي العامِّيّ الذي أنتجه المتكلم وقابله بنموذجه الفصيح المُقارب له تركيبيًّا لندخل بيانات حقيقية لتعليم الآلة الفكر الذي تقدمه هذه النماذج فتحلّل هذه المواد اللغوية وتدرس تركيباتها في خوارزميات بواسطة تقنيات الذكاء الاصطناعي. وقد تطلَّب هذا إجراء عملية واسعة ودقيقة لجمع عينات كافية لجُمل عربية من مستخدمين أردنيين فيها أخطاء لغوية مختلفة أو باللهجة العامية المحلية، ثم كتابة النص اللغوي السليم مقابل كل عينة. ولأن تعلم الآلة العميق يحتاج عينات كثيرة جدًّا غير متوفرة وجمعها مكلف، فقد دعّمنا عمليات تعلم الآلة في البحث بعينات متوفرة أدخلنا عليها أخطاء اصطناعية، إضافة إلى توظيف طرق نقل التعلم لتطوير حلول تعلم الآلة التي تعتمد على هذه العينات وتترجم النصوص المُدخلة إلى نصوص سليمة ومُشكّلة. كذلك يتضمن البحث تطوير تطبيقات للإنترنت والهواتف المحمولة تُضمَّن فيها هذه الحلول لتوفيرها للمستخدم الأردني.

ويظهر في شكل 1 أسلوب الحل المعتمد في هذا المشروع؛ إذ يتضمَّن نموذجي تعلم الآلة، وهما: نموذج التّصحيح، ونموذج التشكيل. يصحّح نموذج التصحيح النصوص العربية المُدخلة باللهجة الأردنية أو التي تحتوي أخطاء لغوية إلى نصوص عربية سليمة؛ ففي المثال نموذج النص المدخل بالعامية هو: "بدي نام" وناتج نموذج التصحيح هو نص فصيح "أريد أن أنام". ومن ثَمَّ يضيف نموذج التشكيل حركات التشكيل (ضبط النص) إلى ناتج نموذج التصحيح لنحصل على النَّص المشكَّل "أُرِيدُ أَن أَنامَ". علمًا بأن استعمال نموذج التشكيل هو اختياري؛ إذ يمكن للمستخدم استعماله والحصول على نصوص مشكلة أو عدم استعماله والاكتفاء بالنص المصحَّح.



شكل1: أسلوب الحل المعتمد في المشروع

أمَّا النَّموذجان الحاسوبيَّان المستعملان في هذا الحل فمشتقان من نموذج لغة كبير مدرَّب مُسبقًا على العديد من اللغات منها العربية، وهو نموذج (ByT5) الذي طوَّره باحثون من شركة جوجل [2]. وهذا النموذج مشتقٌ من نماذج عائلة محوّلات نقل

النص إلى نص (text–to–text transfer transformers, T5)، ويمتاز بتعامله مع النصوص بترميز الأحرف، وليس بترميز الكلمات، لديه قدرة أفضل في التعامل مع النصوص التي تحتوي أخطاء إدخال وأخطاء إملائية. وقد جرى تدريب نموذجي المشروع باستعمال قاعدتي بيانات خاصّتين هما:

1) **قاعدة بيانات اللهجة الأردنية** (JODA) وهي قاعدة البيانات التي جُمعت ضمن هذا المشروع وتتضمن ما مجموعه 59,135 نصًّا أو جملة منها 36,886 نصًّا باللهجة الأردنية و22,249 نصًّا باللغة العربية القياسية الحديثة تتضمن أخطاء إملائية ولغوية متنوعة. وقد جمع هذه القاعدة مساعدو بحث مختصون بعلم البيانات من مصادر متعددة لنصوص أردنية من منصات مثل يوتيوب وفيسبوك وانستغرام وأفلام قصيرة أردنية وقواعد بيانات متخصصة. واضطلع مساعدو بحث مختصون باللغة العربية بإضافة النص العربي الفصيح السليم مقابل كل نص من نصوص هذه القاعدة.
2) **قاعدة بيانات تشكيلة** [3] (Tashkeela–400) وتتكون من 400,000 نصّ من نصوص اللغة العربية الكلاسيكية جُمِعت من الكتب التراثية العربية القديمة واختيرت بعناية على نحو يتضمن فيه كل نص نسبة تشكيل لكل حرف لا تقل عن 80%.

وقد استُخدمت قاعدة بيانات اللهجة الأردنية لتدريب نموذج التصحيح على تحويل النصوص العامية إلى نصوص فصيحة، واستُخدمت قاعدة بيانات تشكيلة لتدريب نموذج التصحيح على تصحيح النصوص التي تحتوي على أخطاء لغوية مختلفة، وتم هذا بإضافة أخطاء اصطناعية آليًّا على هذه القاعدة وتدريب النموذج على إعطاء النص الأصلي السليم من النص ذي الأخطاء. واستخدمت قاعدة بيانات تشكيلة مع نموذج التشكيل لتدريبه على إعطاء النص المشكّل من النص المنزوعةِ منه حركاتُ التشكيل.

3. البُعد اللغوي بين العامية والفصحى

يتعلَّق هذا المحور بإمكانية الوصول إلى حدٍّ يمكن فيه تقريب العامية من الفصيحة، وهذه الإمكانية تنطلق من فرضية أنّ الفصيحة أصلٌ وأنَّ العاميةّ بمختلف لهجاتها متفرعة عنها بمجمل ألفاظها ودلالتها عبر اتصالها بواقعها الثقافي الذي عُبِّر عنه باللغة. وفي الآن ذاته ينطلق من أن نظام التعبير اللغوي يحتاج إلى مستويَي اللغة الفصيح والعامي؛ ذلك أن اللغة بمفهومها العميق والبسيط عند ابن جني هي "أصواتٌ يعبِّرُ بها كلُّ قوم عن أغراضهم"؛ فلم يشترط ابن جني مثلا في مفهومه هذا الفصاحة أو العامية، بل اشترط عناصر صوتية وتعبيرية تنبثق عن واقع القوم اجتماعيًّا؛ إذ إنّ "كل نص هو عبارة عن منظومة لغوية لها قوانينها وآلياتها وترهين لألفاظها في سياقاتها المختلفة" [4]، ممَّا يعني أنّه لا بُد للمتكلم بالعربية أو غيرها من استخدام مستويي التعبير، ولكنه يحتكِم في ذلك إلى المجتمع والمقام والزمان والمكان الذي يُحتاجُ فيه لكل منهما، علاوة على أن العلاقة بين المستويين علاقة تكامل واتصال، لا علاقة تنافر وانفصال.

> *"فاللغة من حيث هي نظام معرضة كي تعيش لفعل الاختراق والتخريب الدائمين ولكن في حدود ما يسمح لها بأن تبقي أرضًا للتفاهم والتخريب لا يلغي مساحة الاصطلاح، ولا يلغي حركة الانتظام في تبدلها، بل هو حياة اللغة وتطورها."* [5]

ونظرًا للواقع الذي نحياه مع اللغة العربية نتيجة الاتصال الثقافي العابر للقارات المتسارع زمنيًّا ومكانيًّا، ونظرًا لكثرة اللهجات العامية التي تفرعت عن المستوى الفصيح، وضعف استخدام اللغة الفصيحة في التواصل مع الآخرين الذي نجم عنه كثير من الأخطاء في الكتابة والصرف والنحو ما أدى إلى اختلاط التعبير وحصول عسر في الفهم، فقد جاء هذا المشروع ليعالج جانبًا من "المحكيات" أو "العاميات" التي ظلت قرونًا طويلة مُهملَةً من علماء اللغة العربية فلم تُدرَس لذاتِها وظل الكلام عليها مرهونًا باللغة الفصيحة وضرورة الحفاظ على نقائها، ولأن المحكيات كذلك عُدّت أنماطًا شفوية فأُهملت آدابُها ولم تحظَ بالتوثيق الملائم الذي يُساعد الدارسين على تتبع حركتها التعاقبية.

وقد كانت تجربتنا في العمل على تحويل النصوص العامية من المحكية الأردنية إلى اللغة الفصيحة واحدة من التجارب الرامية إلى التقعيد لمنهج جديد في دراسة المحكيات؛ فرغم أنه عمل يقوم على تعليم الذكاء الاصطناعي تحويل الجُملة العامية إلى جملة فصيحة، وهو بذلك يُشبه بعض الجهود المبذولة سابقًا في كتب "ردّ العاميّ إلى الفصيح"، فإن المزية الجديدة في هذا العمل هي الاستعانة بالنصوص الحية المأخوذة من قلب اللغة المتداولة بين الناس في مواقع التواصل الاجتماعي؛ فأن تتمرسَ في معالجة آلاف الجُمل التي يتحدّث بها الناس ويكتبونها على مواقع التواصل الاجتماعي هو مقدّمة لمنهج جديد في التعاطي مع المحكيات العربية، يقوم على الاستماع إلى اللغة من الداخل عبر نصوص طبيعية غير مصطنَعَة، **وهذا منهج من شأنه أن يقيم لبنةً أولى في دراسة المحكيّات تقوم على الدّراسة الداخلية للنصوص بُغية الوصول إلى توصيفها توصيفًا شاملًا والتقعيد لها في البنى الصوتية والمعجمية والتركيبية في المستقبل القريب.**

وعليه، فقد ارتأينا أن نطرح جملة من النماذج المستوحاة من النُّصوص العامية الحيّة على كل مستوى من مستويات التداخل آنفة الذِّكر، مع طرح البديل الفصيح لها وملاحظة التغَّيرات فيها.

أ. التداخل الصَّوتي والصَّرفي

إنَّ اللغة في مستواها الفصيح تُعنى ببنية الكلمة صرفيًّا وصوتيًّا وتوليها اهتمامًا بالغًا في صياغة المعنى الذي تؤدِّيه الكلمة خِفَّةً وبلاغة. فتهتم بمباحث القلب والإبدال والحذف والزِّيادات وغيرها، وهي مباحث يتداخل فيها الجانب الصَّرفيّ مع الجانب الصَّوتي. وينطبق الأمرُ ذاته على اللغة بمستواها العامّيّ –اللهجة– فنجد المتكلِّم يلجأ لتقنيَّات صوتية وصرفية ينحرف فيها عن البنية الفصيحة على مستوى المفردة والجملة، إلَّا أنَّ الفارق في مستويي اللغة أن الفصيح يقبل التقعيد والتَّنميط فيبدو من السهل أن تحاكيه الآلة، ولكن العامي لا يجري على قاعدة يمكن اعتمادها في كل آن؛ إذ تتعدَّد اللهجات العامية في البلد الواحد، وتختلف أكثر فأكثر إنْ توسَّعنا إلى مختلف الأقطار العربيَّة؛ لذلك عمدنا في هذا المشروع إلى مقابلة نصوص عامية حيَّة بما يوازيها في النصوص الفصيحة من دون إجراء تغييرات جذرية على النص العامي، بل بالمحافظة على أصله قدر المستطاع. وقد رصدنا جملة من النماذج التي يتداخل فيها المستويان من جهة الصرف والصوت.

أمَّا من الجانب الصَّوتي فمن أبرز الأمثلة على هذا التداخل إبدال ميم الجمع في الضمائر أو هاء الضَّمير واوًا، فمثلًا يُقال بالعامية: " فاقدكو وجودو وعدمو واحد"، ويقابلها بالفصيح: " فاقدكم وجوده وعدمه واحد"؛ فنلحظ أن المتكلم قد أبدل ميم الجمع في الكلمات "فاقدكم، أنتم، عندكم" واوًا تسهيلًا للنُّطق، وكذا فُعِل في هاء الضمير في الكلمات "وجوده، وعدمه". ومن ذلك أيضًا:

الجدول 1: مثال للتداخل في الجانب الصَّوتي بين العامية والفصيحة

النَّصّ العامّيّ	النَّصّ الفصيح
إنتوا ما عندكو ضمير	أنتم ليس عندكم ضمير

ويبيّن الجدول الآتي حالات أخرى من التَّداخل في الجانب الصَّوتي بين العامية والفصيحة مع ملاحظة التَّغيُّرات عليها من خلال الأمثلة والشرح:

الجدول 2: حالات أخرى من التَّداخل في الجانب الصَّوتي بين العامية والفصيحة

التعبير العامّي	التعبير الفصيح	الشرح
حالة (1) قلب الحروف وإبدالها		
لسا لسا بدهم وقت	ليس بعد، ليس بعد، يحتاجون وقتًا	فكلمة (لِسَّا) أصلها ليس، فقلب الحروف (لِسيّ) ثم أبدل الياء ألفا لأنَّ ما قبلها مفتوح فناسبته الألف ليكون أخفّ في النطق.
قلة الدين والكذب اشي عادي لهيك وصلنا لهون	قلة الدين والكذب شيء طبيعي لهذا وصلنا إلى هنا	كلمة (هون) أصلها هنا، فقلب مكان الألف ثم أبدلها واوا، ولها طرق أخرى للفظ حسب اللهجة العامّية فقد يقال: هون، هنيك، هان.
حالة (2) إضافة حرف للكلمة		
بالنهاية شغل ضياء الو مش للناس	في النهاية عمل ضياء له لا للناس	كلمة (الو) يقابلها (له) في الفصيح، لكن في العامّيَّة يُضاف همزة قبل حرف الجر لِ المتصل بالضمير له = اِله، إلها تسهيلًا للنُّطق.
مين هو؟	من هو؟	في كلمة (من) أضاف الياء إلى اسم الاستفهام مَن، لعلة صوتيَّة.
حالة (3) تغيير حرف كامل في الكلمة لسهولة النطق		
يا ريت والله	يا ليت والله	في كلمة (ليت) أبدل اللام راءً في حرف التمني ليت لتقاربهما مخرجًا فاحتملت هذا التغيير.
وينك إنت؟	أين أنت؟	في كلمة (أين) أبدل الهمزة واوًا لسهولتها في النُّطق.
حالة (4) حذف حروف من الكلمة		
هاي البنت الي راكبه سياره الي بصف تاسع بحب أخبرك انو اللي اشتراها أبوها من عرق جبينه	هذه البنت التي تركب سيارة، التي في الصف التاسع، أحب أن أخبرك أن الذي اشتراها أبوها من عرق جبينه	فاسم الإشارة (هذه) والاسم الموصول (الذي) و(التي) في العامية حذف منهما حرفي (الذال) و(التاء)، فيقال في العامية هاي البنت أو هالبنت، والّي، لتسهيل النطق.
حالة (5) تغيير حركات الكلمة		
كِيفْ علاماتك؟	كَيفَ علاماتُك؟	فاختلفت طريقة نطقها ولم تختلف صيغتها.
حالة (6) حذف حروف من التراكيب		

ول عليك مثقل دمك بينك وبين التمثيل بلاد	عجبا لك! ما أثقل دمك! بينك وبين التمثيل بلاد	في أسلوب التعجب يحذف همزة أفعل التفضيل، فيقول في العامية "مثقل دمك" ومثلها: مبخلك = ما أبخلك مكرمك = ما أكرمك محنك = ما أحنك وهكذا.
ثمر فيك اخوي؟	أثمر فيك أخي؟	في أسلوب الاستفهام كثيرًا ما تحذف أداة الاستفهام الهمزة أو هل اعتمادًا على السياق التداولي للخطاب أو على علامة الترقيم (؟)، ويظهر ذلك في اللفظ الصوتي الذي يلفظه المتكلم.

وأمَّا من الجانب الصرفي فإن من أبرز الأمثلة على ذلك في العامية ما يجري من تحويرات على بنية الفعل بأزمنته الثلاثة الماضي والمضارع والأمر؛ فالماضي الفصيح للفعل "صَحوت" يقابله في العامِيَّة قولنا: اِصْحيت، أو صْحيتْ، أو أن نعبِّر عنه باستخدام صيغة الفاعل، صاحِي. والمضارع الفصيح أصحو، في العامِيَّة بَصْحَى، فاختلف حرف المضارعة من الهمزة إلى الباء، ومثلها ببيع، بشرب، بنام ... إلخ، وقد يكتسي التَّعبير عن المضارع بالعامية بعدًا أكثر غرابة حين يستعين المتكلِّم بصيغة اسم الفاعل قبل الفعل المضارع وهو ما نسميه بالفعل المساعد فيقول: "قاعد بصحى" حيث أضاف التعبير (قاعد + ب) قبل الفعل لوصف حالة فعله. ومن ذلك أيضًا: "أنا قاعدة بموت" وتفصيحها "أنا أموت". وفي صيغة الأمر تارة يحافظ على بنائه الفصيح في العامية مثل "اِصحَ"، وتارة يجمع في وقفه بين ساكنين فيلتقيان، وهذا من المحظور أن يقع في الفصحى مثل: قُول، موت، نامْ ...

هذا تمثيلٌ عام على بعض الحالات، ونوصي بإجراء دراسات موسَّعة من شأنها أن تغني هذا الجانب في الدراسات الصوتية والصرفية للهجات المحكيَّة؛ لملاحظة الاختلافات في هيكل بنية الكلمة من جهة الصّرف والصَّوت نتيجة الإبدالات والتغييرات الصوتية والإملائية.

ب. التداخل التركيبي النحوي

يمكن إجمالًا القول إنّ ثمَّة اتّفاقًا عامًّا بين تركيب الجملة في العامية والفصيحة، من جهة علاقات الإسناد، إلا أنّ الترتيب الذي تقدِّمه العامية قد يُشوِّش قليلًا إما لضعف المتكلم وعجزه عن توظيف أفكاره بقوالب عامية مناسبة، وإما لاعتماده على كثرة الحذف في التعبير اتكاءً على سياق التداول؛ فيظهر التركيب كما لو أنَّه قاصر عن إيصال الفكرة ولا سيما إذا ما قُدّمَ مكتوبًا، وهو ما قد تواجه الآلة صعوبة ما في مقابلته بمرادفه الفصيح على مستوى التركيب. وينجم عن هذا التشويش تباين بين النصوص العامية المكتوبة في جودة التعبير وصحة انتقاء المفردات وحُسن تنظيم بناء الجملة، وقد كان مؤسفًا أن عددًا كبيرًا من العينات التي حوّلناها إلى الفصيحة كان يفتقد إلى خصائص التعبير الصحيح؛ كأن تتداخل عناصر الإسناد في الجملة (مثال: مبتدأ خبرُه غير موجود، أسماء موصولة في غير مكانها، سقوط أل التعريف من الكتابة، استخدام المتكلم للضمائر فلا يسندها لمرادفها كما هو الحال في الفصيحة، أو في أزمنة الأفعال، أو في النَّفي والإيجاب، أو في أسلوبي التعجب والاستفهام.) فضلًا عن مشكلة عدم القدرة على محاكاة الإملاء الفصيح. ونلخصها بالنماذج الآتية.

النموذج الأول: الاضطراب اللغوي

"هو قلبه نضيف رغب الاشياء الموجعه الي حصلت له ركز للايجابي وهذا اكبر دليل انك طبق كلام الله واستحق النجاح" فهذا التعليق يمدح المتابِع من خلاله الشخصية التي يتابعُها بأنه ذو قلب نظيف ورغم الأشياء الموجعة التي حصلت معه فإنه تجاوزها وركز على الإيجابيات تطبيقًا لكلام الله، ومن ثَمّ فقد استحق النجاح.

ولكن يتضح أن هذا المعنى قد تعرض في العبارة التي كتبها المعلق لتعقيد كبير وما يمكن تسميته بالتعبير المختنِق؛ ففضلًا عن عدم مراعاة الكتابة الإملائية الصحيحة، فلا يوجد ربط لأجزاء الكلام، كذلك لا يوجد استخدام صحيح لحروف الجر مع الأفعال، وفضلًا عن ذلك هناك اضطراب واضح في الأفكار.

ويجدر التَّنبيه إلى أنَّ ذلك يظهر من وجهة النظر التي تطرحها اللغة الفصيحة، أمَّا لو نظرنا لها من وجهة نظر اللهجة العامية فإنَّ هذه الإشكالات لا يُعبأ بها؛ ذلك أنَّ المعرفة التي يريد المتكلم أن تصل للمخاطب في مجتمعه قد وصلت، حتى إنَّه لم يحتج لكثير من أدوات الربط التي لا بد من إضافتها إلى الجملة في مقابلها الفصيح. وهذا ما يجعل من المحاولة التي يقدمها المشروع في معالجة هذه الهُوَّة بين العامية والفصيحة أمرًا قد يحل إشكالًا كبيرًا، لا سيَّما فيما سيقدمه في مجالات الترجمة وتعلم العربية لغير ناطقيها.

وهذا مثال آخر يمكن للقارئ أن يكتشف قدر ما فيه من اضطراب تعبيري:

"انا من وجهة نظري انت انسان عندي مباااادي واصول والزمان والمكان لايغيرك الا لافضل شخص وعندك عائلتك هم الاهم ومو ناسي اهل بلدك الاردن ربنا يمتعك بالصحة والعافية ويوفقك دنيا واخرة مشكورة انك كان عندك ضيف بمستوي الفهم هذا"

يمكن أن يقول القارئ إن هذه الجملة تبدو واضحة من جهة التعبير العامِّيّ، إلَّا أنَّه يعتريها كثيرٌ من الاضطراب بسبب الأخطاء الإملائيّة أو الطِّباعيّة كما في قوله (عندي، مباااادي، لأفضل، بمستوي، ...)؛ فالتركيب العامي هنا حافظ على ترتيب الجملة المعتاد في الفصيحة، ولكنه خالف في طريقة استخدام الضمائر. وقد عمدنا لوضع مقابلها الفصيح بُغية تدريب الآلة على هذا النمط من التعقيد في التركيب ليشتغل عليها من جانب التركيب والإملاء والأخطاء الطباعية.

النموذج الثاني: استخدام القوالب الجاهزة

تستخدم العامية **بعض المفردات التي أشبه ما تكون بقوالب جاهزة لا يُعرف أصلها وإنما أنتجتها الثقافة فأصبحت دارجة في الاستخدام اليومي**، وغالبًا ما تكون هذه المفردات عوضا من تركيب كامل يحتوي مسندًا ومسندًا إليه في الفصيحة. ومن أبرز الأمثلة على تلك العبارات:

الجدول 3: أمثلة من العامية التي تستخدم مفردات بقوالب جاهزة

التعبير العامِّي	التعبير الفصيح	الشرح
بديش الثقه، بدي أموت	لا أريد الثقة، أريد أن أموت	كلمة (بِدّي) و (بِدّيش) في العامِّية يرادفها في الفصيح تركيبٌ مختلف عنها تمامًا وهو "أريد" و "لا أريد".
مهو ما صنعهن غير الانسان	لكنّه لم يصنعهم غير الإنسان	كلمة "مهو" تُفصَّح بـ "لكنّه" وهي كلمات تختلف عن العامِّية.
طول بالك خالي	اصبر خالي	طَول بالك، جملة عامِّية دارجة تعني اصبر أو تريَّث أو انتظر.
سُولف معي	تكلَّم معي	سُولِف جملة عامّيَّة تعني تحدث معي أو تكلم، ومنه قولهم "سوالف حصيدة"، أي أحاديث لا قيمة لها.
الطلاب بلَّشوا الامتحان	الطلاب بدأوا الامتحان	كلمة "بلَّش" و"بلشان"، بمعنى بدأ، بادئ، وهي جملة كاملة تحتوي مسندا إليه.
هوة كمان مخلص	هو أيضًا مخلص	كلمة "كمان" بمعنى أيضًا.
عشان ما ارسب من وراكم	كي لا أرسب بسببكم	كلمة "عشان" أو "مشان" بمعنى كي أو لأجل بالفصحى.
إذن كلهم متل بعض وما تشد ع معدتك كرمال حدا	إذن كلهم مثل بعض ولا تشد على معدتك لأجل أحد	كلمة "كرمال" بمعنى لأجل بالفصحى.
بس الوضع صعب خليها ع الله بس	لكن الوضع صعب اتركها لله فقط	كلمة "بس" في العامية تختلف باختلاف السياق الذي ترد فيه ففي مطلع هذا التركيب تعني لكن، وفي نهايته تعني فقط، وقد تتفرع إلى معان أخرى بحسب السياق.
عمل زي الي قبله	عمل مثل الذي قبله	كلمة زي وتستخدم للتشبيه في العامية بمعنى مِثل.
هو أحسن ولَّا صديقي المسيحي الي وقف معي وقفة رجال؟	هو أحسن أم صديقي المسيحي الذي وقف معي وقفة رجال؟	كلمة "ولَّا" التي يقابلها في الفصحى "أمْ".

النموذج الثالث: أسلوب الاستفهام

تركيب صيغ الاستفهام في العامية والفصيحة متنوّع؛ إذ قد تتّفق في ألفاظها تارة وقد تختلف إمَّا كلِّيًا أو باختلاف الصوت والصِّيغة. فمثلًا: الاستفهام بكلمة "شو" و"شِنو" و " إيش" التي تفصَّح بـ ما؟ أو ماذا؟، فهنا نلحظ الاختلاف الكلي بين صيغة العامية والفصيحة، ومن ذلك قولهم: "وصاحب العقار شو وضعه؟" = تفصيحها: وصاحب العقار ما وضعه؟

وكذا الاستفهام بـ" ليش" التي يقابلها في الفصيح "لماذا؟"، مثلًا: ليش رُحت ع الجامعة؟ وتفصيحها: لماذا ذهبت إلى الجامعة؟

النموذج الرابع: أسلوب النَّفي

هناك تراكيب خاصة للنفي بالعامية تختلف عن النفي بالفصيحة، من أشهرها: مِش، ما عِوضًا عن لم، بديش، بدوش، إضافة الشين في نهاية الفعل المضارع. ومن الأمثلة على ذلك:

الجدول 4: أمثلة تراكيب خاصة بالنفي

التعبير العامِّي	التعبير الفصيح	الشَّرح
والله ما حبيتك وما تقبَّلتك	والله لم أحبك ولم أتقبلك	ما = لم
هاذ الطَّالب مجتهد ونشيط بس مش هادي	هذا الطَّالب مجتهد ونشيط لكنّه غير هادئ	النفي بكلمة مِش ويقابلها في الفصيح "غير".
الشَّب اللي بتركش الكتاب من ايدو "بس والله بيستحوش بالمرة	الشاب الذي لا يترك الكتاب من يده لكن والله لا يستحون نهائيًّا	بتركش= لا يترك إضافة الشين في النهاية بدل لا النافية في التعبير الفصيح بستحوش= لا يستحون

ج. التداخل الدلالي

ونعنى بالتداخل الدلالي ورود كلمة واحدة بمعانٍ مختلفة بحسب السياق الذي تدُلّ عليه، ككلمة (بس)، و(بعدين).

- فمثلًا في جملة (باكل بس معك) أتت كلمة (بس) بمعنى (فقط) وتفصيحها (آكل فقط معك)، ولكن في سياق آخر يقول فيه المتكلم: (أنا بحترمه بس هوة ما بحترمني)، فقد جاءت كلمة (بس) بمعنى (لكن)، وتفصيحها (أنا أحترمه لكنه لا يحترمني)، وفي سياق آخر يقول فيه المتكلم (بس ولا كلمة) جاءت كلمة (بس) بمعنى (كفى) أو (اصمت). فنلحظ أنَّ الكلمة الواحدة لها عدة احتمالات على مستوى الدلالة تحدد بحسب السياق.
- كلمة (بعدين) وتفصيحها (بعد هذا)، وقد تفصح بمعنى ظرف الزمان (لاحقًا)، وقد تدل على فعل لا يتعلق بالزمان تمامًا كقول المتكلم: " بعدين معك"، ومعناه التذمّر والغضب واستنفاد الطاقة مع المُخاطَب، ولا شكّ أن ثمة علاقة زمانية موجودة في هذا الاستخدام تتعلّق بطول وقت الصّبر على المُخاطَب، ولكن لا يعني تعبير "بعدين معك" الظرفية الزمنية تمامًا.

د. التداخل البلاغي

نقصد بهذا البعد بحث القضايا البلاغية التي ينتجها متكلم العامية سواء في الأمثال الشعبية أو المسكوكات اللفظية التي تتضمن كنايات واستعارات وغير ذلك؛ ذلك أنَّ **البعد البلاغي في العامية لافتٌ للنَّظر؛ فهناك كنايات واستعارات تمثِّل نتاجًا ثقافيًّا يربط بين العامي والفصيح والوسط الذي تشكَّلت فيه**، بعضها مأخوذ من الأمثال الفصيحة وبعضها أنتجته الثقافة الخاصة بهذه اللهجة العامية، وبعضها مُشترك عابر للثقافة العربية عمومًا.

وقد استقرأنا أشهر الأمثلة من النّماذج التي اشتغل عليها المشروع في عمليَّة الحوسبة، ملتزمين قدر الإمكان في عملية تعليم الآلة بالحفاظ على التَّركيب كما هو في غالب الحالات، وتعليم الآلة هذه التراكيب من شأنه أن يُقعِّد للواقع الثقافي الذي تشغله هذه العامّية.

من ذلك قولهم في العامية: "منسمع قصص بتشيب الرأس"، وهذه استعارة مستخدمة في الفصحى، تعبيرًا عن صعوبة الأمر، وتفصيحها: "نسمع قصصًا تشيب الرأس"، ولا شكّ أن هذا التعبير نفسه موجود في الأدبيات الفصيحة، وقد ورد أيضًا في القرآن الكريم: "فَكَيْفَ تَتَّقُونَ إِنْ كَفَرْتُمْ يَوْمًا يَجْعَلُ الْوِلْدَانَ شِيبًا" (المزمّل: 17).

وكذا قولهم في العامية: "جابهم اكتاف والله"، وهذه كناية عن القوة الهائلة للشخص، وتفصيحها "قضى عليهم"، ولعدم شهرة هذا التعبير فقد ارتأينا تحويله إلى الفعل قضى مع الاستغناء عن حقيقة التركيب، وهذا الإجراء لجأنا إليه في حالات قليلة. ويبيّن الجدول الآتي أمثلة أخرى للتعبير العامّيّ ومقابله الفصيح والمعنى الذي تشكله الكناية أو الاستعارة في التعبير.

الجدول 5: أمثلة التداخل البلاغي

التركيب العامّيّ	التركيب الفصيح	المعنى الكنائي أو الاستعاري
متنفس لحاله وعن حاله شغل بهلوانات ولا فيهم من يمثل شعره من راس مواطن مجلس ال ٢٩ %	متنفس وحده وعن نفسه، شغل بهلوانات، وليس فيهم من يمثل شعرة من رأس مواطن، مجلس الـ ٢٩ %	شغل بهلوانات كناية عن الاستعراض الذي لا طائل منه. لا يمثل شعرة كناية عن اللاجدوى فيما يفيده.
لا لسا ما اكلت هوا	لا ليس بعد، لم آكل هواء	استعار الأكل للهواء، كناية عن التعب وسوء الحظ.
هيك بتضرب السياحه	هكذا تضرُّ السياحة	استعارة الضَّرب في العامية للدلالة على وقوع الضرر. وقد حوَّلناها في التَّفصيح إلى معناها الحقيقي الضَّرر لتكون أقرب للفهم.
المدارس الخاصة فيها تمييز اللي بتجيب هدايا للمعلمات وتمسح جوخ للاداره بصير ابنها او بنتها يتركز عليه	المدارس الخاصة فيها تمييز، التي تحضر هدايا للمعلمات وتمسح جوخًا للإدارة يصبح التركيز على ابنها أو ابنتها	مسح الجوخ كناية عن التَّملُّق.
الطلاب والاهل عباره عن حقل تجارب كل واحده يستلم هالوزاره بدو يفرض عضلاته على الطلاب	الطلاب والأهل حقل تجارب، كل شخص يستلم هذه الوزارة يريد أن يفرض عضلاته على الطلاب	فرض العضلات كناية عن القوة في غير مكانها.
نحن ما منحكي على الشرطه عنا الشرطي لما يشوف وزير ولا منصب معين صعب تخالفه لانه بخاف يروح ورا الشمس شفت الفرق	نحن لا نتكلم عن الشرطة، عندنا الشرطي حين يرى وزيرا أو منصبا معينا فمن الصعب أن يخالفه لأنه يخاف أن يذهب وراء الشمس، رأيت الفرق؟	لأنه يخاف أن يذهب وراء الشمس، كناية عن النهاية الوخيمة إن فعل أمرًا ما. وهذا تعبير مستخدم بكثرة في الفصحى.
لازم تشديد وعقوبات صارمة خللي السير بالبلد يصير متل العالم والناس	يجب التشديد مع عقوبات صارمة ليصبح السير في البلد مثل العالم والناس	تشبيه شيء معين بقولنا "مثل العالم والناس" أو "مثل الخلق". كناية عن الجودة والشيء الحسن.
انا اكلتها يومها ومشوار بالصيف وازمة يعني استغربت	أنا تأذيت يومها ومشوار في الصيف وازدحام، أعني	التعبير بالفعل أكلها يكثر في العامّيَّة كناية عن التّعرُّض لأذى أو تحمُّل الشيء عنوة.

شو هل انسان ما بهمو صحة الي معه	استغربت، ما هذا الإنسان! لا تهمه صحة الذين معه!	وفي التَّقصيح حوَّلناها لمعناها الذي تكنّي عنها إيثارًا لتوضيح المعنى.
اكلونا و اكلو حقوقنا و احنا بنركض ورى ليره و ليرتين	أكلونا وأكلوا حقوقنا ونحن نركض وراء ليرة وليرتين	أكلني بمعنى أضاع حقي، وهذا ينطبق عليه ما عرضناه في المثال السَّابق.
بتحكي كلمه كويسة أو تبلع لسانك	فإما أن تقول كلمة جيدة أو تبلع لسانك	تبلع لسانك كناية عن ضرورة لزوم الصمت في حال عدم القدرة على الكلام الجيد، أو عدم الرغبة في الكلام.
ياخي الممثل الاردني يطلع سلم طلقه وينزل كرفته	يا أخي، الممثل الأردني يصعد السلم بسرعة وينزله دحرجة	كناية عن التراجع السريع في المستوى بعد ارتفاع كبير.
أبدعت يكفي انك تشكو هموم المواطنين مش زؤ زملائك نايمين نوم أهل الكهف	أبدعت، يكفي أنك تشكو هموم المواطنين، لا كزملائك النائمين نومة أهل الكهف	كناية عن الكسل وعدم العمل فشبههم بأهل الكهف.
مافي غبرة ع كلامك!!	ما من غبار على كلامك	كناية عن جودة الكلام.
أطلع احكي على الحكومة وانتقدها وشد على حالك وبعدين الموازنة وغيرها بمشي زي الحلاوة	تخرج تتكلم عن الحكومة وتنتقدها وتتعب نفسك، وبعد ذلك الموازنة وغيرها تمشي مثل الحلاوة	تمشي مثل الحلاوة كناية عن سهولة الأمر.
اذا رئيس الوزراء ما بيسمع كلامك على قولت المثل على بال مين يللي بترقص بالعتمة	إذا رئيس الوزراء لم يسمع كلامك، على قول المثل: على بال من يا من ترقص في العتمة	"على بال من يا من ترقص في العتمة " هذا مثلٌ يُضرَب للشخص الذي لا يُكترَث لِما حلَّ به.
عند عمك طحنا	عند عمك طحنا	كناية عن ذهاب الجهد بلا طائل، أو الاستخفاف بالجهد المبذول لعمل ما.
حط بالخرج الحصيدة حفظناها وابر البنج بطلت تعطي مفعول.	ضع في الخرج، الحصيدة حفظناها، وإبر التخدير لم تعد تعطي مفعولا	حط بالخرج مثل يكنى عنه بعدم الاكتراث بالشيء لطول المكث والانتظار. إبر البنج استعارها كناية عن التّسكين وانتظار الأمر حتى يحصل.
خض المي وهيه مي	خض الماء وهو ماء	كناية عن عدم التغيير.
نفس نائب من عشيرتي لا بهش ولا بنش	نفس نائب من عشيرتي لا في العير ولا في النفير	كناية عن قلَّة نفع هذا الشخص فلا ينفع في الأمور البسيطة ولا المعقدة. وقد قابلناه بما يرادفه في الفصحى "لا في العير ولا في النَّفير".
كلام ذهب بس بتفخ بقربه مخزوقه	كلام من ذهب، لكنك تنفخ في قربة مخروقة	القِربة المخروقة كناية عن فقدان الأمل.

كلهم ما حد بيقلي بيضه	كلهم لا أحد فيهم يقلي بيضة	كناية عن عدم القدرة على فعل الشيء أو التأثير فيه.
احنا يلي اعطناهم عين كان كل ما يعتدوا على الشعب والمسجد انضل ساكتين وما نسوي اشي	نحن الذين أعطيناهم عينا، كنا كلما اعتدوا على الشعب والمسجد نبقى ساكتين ولا نفعل شيئا	أعطيناهم عينًا كناية عن تمكينهم واحترامهم ثم لا طائل من هذا الإكرام.
علاماتك هذه اغسلها واشرب ميتها	علاماتك هذه اغسلها واشرب ماءها	كناية عن عدم تحصيل النفع من الأمر المتحقِّق.
لا تشدي ع حالك ليطقلك عرق	لا تشدي على نفسك خشية أن يطق لك عرق	يطق لك عرق كناية عن الغضب الشديد.
هه لاااا مش مصدقة كتر خيركم كنا مستنينهم على نار!	ههه، لا، لا أصدق، كثر خيركم، كنا ننتظرهم على نار!	كناية عن الاحتفاء الشديد بالأمر المنتظر.
الكل جاي حبه عالطاحون	الكل سيأتي حبه على الطاحون	كناية عن التهديد بالإيذاء أو ما شابهه.
طالما وضعكم بأوروبا لووز	ما دام وضعكم في أوروبا لوزا	وضعكم لوز كناية عن جودة الوضع وحسن الحال.
شو مثلا فيك ريشه فوق راسك بنت معلمه	ماذا مثلا؟ عليك ريشة فوق رأسك، ابنة معلمة	ريشة فوق رأسك كناية عن التميز عن الغير.
انتبهي منيح احسن ما تفوتي بل حيط	انتبهي جيدا كي لا تدخلي في الحائط	فات بالحيط، كناية عن عدم القدرة على تبيُّن الأمر على حقيقته.
من يوم ما راحت مراقبة الاسعار لوزارة التموين نفشو التجار ريشهم وصارو حيتان والمواطن سمكه صغيره زلطوه	منذ أن توقفت مراقبة الأسعار لوزارة التموين نفش التجار ريشهم وصاروا حيتانا والمواطن سمكة صغيرة أكلوها	نفش ريشه كناية عن التكبُّر.
المهم الحكومة تعيش والحيتان تنتعش	المهم أن تعيش الحكومة وتنتعش الحيتان	استعارة الحيتان للتجار أو أصحاب المناصب عمومًا كناية عن ضخامتهم وعدم القدرة على مجابهتهم.
النص الثاني ديون وتلبيس طواقي	النصف الثاني ديون وتلبيس طواق	تلبيس طواق كناية عن تحمل الأمر عنوة.
الكل ماكل روح الخل	الكل يأكل روح الخل	أكل روح الخل كناية عن تعرض الشخص لمصائب كبيرة أو ما شابه.
يا ترى فكرتوا بإعادة هيكلة الشركة وتغير القائمين عليها بدل من ضخ المال على شركة إدارتها واقعه من طياره؟	يا ترى، هل فكرتم بإعادة هيكلة الشركة وتغيير القائمين عليها بدلا من ضخ المال على شركة إدارتها تقع من طائرة؟	وقع من طائرة كناية عن سوء الحال والأوضاع.
أحمد دقّ ع صدره	أحمد دقّ على صدره	دق على صدره كناية عن تكفله بالأمر.

4. الختام والتوصيات

ختامًا، نقدِّمُ بعض النتائج والتوصيات من واقع الاحتكاك بالجملة العامية وتحويلها إلى النموذج الفصيح، وهي نتائج تركّز على الجانب التعليمي والتربويّ الذي قد يُستفاد منه في هذا السياق، بالإضافة إلى بعض النتائج ذات الصِّلة بالجانب الحاسوبي وتعلم الآلة.

1) يُلحَظ أنّ لغة التعليق المكتوب على النصوص المرئية والمسموعة تتباين بتباين الموضوع والمخاطِب بالفيديو؛ فإذا كان الموضوع دينيًّا أو ثقافيًّا أو كان المتحدث يميل إلى "تفصيح عباراته"، فإن استجابة الناس لمحاكاة النموذج الذي يسمعونه تكون أعلى، ومن ثَمّ نجد أن التعبير معتنًى به أكثر وأن الجمل أكثر وضوحًا واتّساقًا، بعكس المحتوى المرئي العادي الذي يتعلق بالمشكلات الأسرية أو بالطبخ أو بنصائح التنمية البشرية، فإن الناس يتفاعلون معه بجودة كتابية أقل ومن ثَم تظهر مشكلات في التعبير.

والتوصية التي نعطيها هُنا أن يُستفادَ من المرئيات التي تقدم محتويات دينية أو ثقافية بلغة تمزج الفصيح بالعامي بُغية تطوير أساليب التعبير لدى الطلبة في مراحل مبكرة على نحو يجعلهم على مقربة من التعبير الفصيح أو على الأقل لأن يكونوا قادرين على التعبير الكتابي العامي الجيد الذي يخلو من مشكلات التعبير المألوفة.

2) يُلحَظ أن نسبة التعليقات باللهجات الخاصة جدًّا التي تختلف من مدينة أردنية إلى أخرى، بل من قرية إلى أخرى هي نسبة قليلة جدًّا وتكاد تكون معدومة؛ فاللهجة الطاغية على التعبير الكتابي وعلى المرئيات التي فرغناها هي ما يُصطلح عليه باللهجة العمّانية التي مزجت منذ مراحل مبكرة بين اللهجات المختلفة من مناطق الأردن وفلسطين، وهذا يعني أن اللهجة العمّانية قد خطت خطوات جيدة في أن تكون لغة عامية معبّرة عن الأردن بوصفه قُطرًا عربيًّا.

ويمكن أن نعطي توصية بهذا الخصوص تتعلق بإعداد مزيد من الدراسات الصوتية والمعجمية والتركيبية التي تقعّد لهذه اللهجة وتدرسُها من عدة جوانب؛ أوّلًا من حيث درجة اقترابها من النموذج الفصيح، وثانيًا من حيث درجة إسهام اللهجات المختلفة في تكوينها وجعلها لهجةً مصطلحًا عليها لدى الجيل الجديد. ومن شأن هذا التقعيد أن يُصدّر اللهجة الأردنية كما استقامت على ألسنة أهلها بعد مئة عام من تأسيس الدولة الأردنية.

3) كان العمل على تحويل الجمل من العامية إلى الفصيحة يتوخى الحفاظ على بنية التركيب العاميّ والابتعاد عن إعادة الصياغة التي تهشّم أجزاء الجملة العامية، مثلًا: "انتا بدك تروح على العرس يوم السبت الجاي؟"، التفصيح يكون: "هل أنت تريد أن تذهب إلى العرس يوم السبت القادم؟"، علمًا أنّ هذه الجملة قد تُحرّر وتُختصَر لو كانت الغاية من العمل هي الوصول إلى جمل عربيّة عالية فتكون: "هل تريد أن تذهب إلى العرس السبتَ القادم؟"، أو جملة "شو عندك جايبني على هالصبح!"، فيصير التفصيح: "ماذا لديك لتأتيني منذ الصباح"؟ مع أنه يمكن أن تكون الجملة أعلى على هذا النحو: "ما وراءك لتأتيني هذا الصباح؟". ومع إمكانية تفصيح الجملة أكثر وتحريرها وجعلها أوقع وأكثر اختصارًا، فقد حرصنا على محاكاة التركيب العامي والاقتصار على تحويل المفردات من دون المساس ببنية التركيب؛ لأن من شأن ذلك أن يعطي صورة عن طبيعة ما يعتري اللهجة المحكية من ظواهر تركيبية تتعلق بالتطويل والتكرار والحشو، وهي من صفات الأقوال الشفوية عمومًا.

4) كان من الطبيعي أن نجد تراكيبَ وجملًا في العامية غير موجودة في الفصيحة، ورغم أنها قليلة فإنها تعطي صورة عن الشفوية العامية التي يصعُبُ كثيرًا إيجاد بديل فصيح مطابق لها؛ مثلًا: "قال شغل قال! البطالة معبية البلد"؛ فتكرار "قال" هنا بين كلمتين لغاية السخرية أو الحَطّ من شأن فكرةٍ ما هو أسلوب عاميّ محض لا نجد له شبيهًا في الفصيحة من حيثُ الاعتماد على مفردة "قال" المكررة أو غير المكررة للتعبير عن معنى السخرية؛ لذلك

اضطررنا إلى تفصيح التركيب ليكون: "أي شغلٍ هذا! البطالة تملأ البلد"، ومن الواضح أن ثمة فرقًا تركيبيًّا بين الجملة الفصيحة والعامية، ومن شأن هذا أن يفتح المجال لدراسة بعض التراكيب العامية وكيفية تطورها لتعبر عن معاني غير موجودة في الفصيحة.

5) لقد سهّل اعتماد المشروع على نماذج لغة كبيرة مدرَّبة مسبقًا من عملية تطوير نموذجي التصحيح والتشكيل، فاستعمل هذا المشروع نماذج لغة كبيرة واستفاد من دِقَّتها دون الحاجة إلى مَوارد حاسوبيَّة ضخمة لتدريبها. لذا نوصي بأن تعتمد حلول حوسبة اللغة العربية –ما أمكن– على النماذج العالميَّة المدرَّبة مسبقًا وإعادة تدريبها وتخصيصها للحلول المطلوبة.

6) لقد كان الجهد الأكبر في هذا المشروع منصبًّا على جمع قاعدة بيانات اللهجة الأردنية (JODA) وتطويرها، فنوصي بشأن تطوير مثل حلول هذا المشروع للَّهجات العربية الأخرى أن تُجمع مثل هذه القاعدة لباقي اللهجات العربية ليصبح من الممكن حوسبتها وفق تقنيات تعلُّم الآلة.

5. المراجع

[1]

رنا الدقاق، نورة الحبسي. اللسانيات الحاسوبية... البدايات والتطبيقات والتحديات. متوفر إلكترونيًا، 26 حزيران 2024، تم الوصول إليه آخر مرة بتاريخ 11 تشرين الثاني 2024.

[2]

Linting Xue, Aditya Barua, Noah Constant, Rami Al-Rfou, Sharan Narang, Mihir Kale, Adam Roberts, and Colin Raffel. ByT5: Towards a token-free future with pre-trained byte-to-byte models, Transactions of the Association for Computational Linguistics, 10, 291–306, 2022.

[3]

Asma Abdel Karim, G. Abandah. On the training of deep neural networks for automatic Arabic-text diacritization. International Journal of Advanced Computer Science and Applications, 12(8), 2021.

[4]

حسين خمري. نظرية النص من بنية المعنى إلى سيميائية الدال. منشورات الاختلاف، ط 1، 2007، ص 266.

[5]

يمنى العيد. في معرفة النص. منشورات دار الآفاق الجديدة، ط 1، لبنان، 1983م.

6. السيرة الذاتية للباحثين

	دلال إبراهيم الزهيري، حاصلة على درجة البكالوريوس عام (٢٠١٧) والماجستير عام (٢٠٢٢) في اللغة العربية وآدابها من الجامعة الأردنية. تعمل حاليًا بصفة محاضر غير متفرغ في الجامعة الأردنية وجامعة اليرموك، اهتماماتها البحثية في مجال اللغة والدَّلالة والتَّأويل.
	حسين فتحي عدوان حاصل على درجة الماجستير في اللغة العربية وآدابها من الجامعة الأردنية عام 2019، ويحضر لدرجة الدكتوراه في الجامعة الأردنية، ويعمل مدققا لغويا في مجال الإعلام (التلفزيون العربي)، وباحث في الدراسات اللغوية والأدبية.
	الدكتور غيث علي عبندة هو أستاذ هندسة الحاسوب في الجامعة الأردنية، حاصل على درجتي الدكتوراة والماجستير في علوم وهندسة الحاسوب من جامعة ميشيغان (1998 و1995) ودرجة البكالوريوس في الهندسة الكهربائية من الجامعة الأردنية (1985). لديه خبرة أكاديمية لمدة 26 عامًا في تدريس وتطوير برامج ومواد في هندسة وعلوم الحاسوب وإجراء البحوث المتعلقة بالذكاء الاصطناعي لخدمة اللغة العربية والتصويت الإلكتروني والحواسيب المتوازية.
	الدكتور يوسف حسين حمدان حاصل على درجة الدكتوراة في الأدب الحديث من جامعة إدنبرة عام 2013. وهو يعمل حالياً أستاذاً مشاركاً في قسم اللغة العربية وآدابها في الجامعة الأردنية.

7. الخلاصة باللغة الانجليزية

Converting Colloquial to Classical Arabic as within the Project Developing Applications to Correct Jordanian Spoken Arabic to Proper Language Using Machine Learning Techniques

Dalal I. Alzuheiri[1], Hussein F. Adwan[2], Gheith A. Abandah[3], and Yousef H. Hamdan[4]
[1,2,4]Department of Arabic Language and Literature
[3]Computer Engineering Department
The University of Jordan, Amman, Jordaan
[1] dalalalzuheiri@gmail.com, [2] husseinadwan71@gmail.com, [3] abandah@ju.edu.jo,
[4] y.hamdan@ju.edu.jo

Abstract
This research explores issues in transforming spoken or written colloquial dialects into standard Arabic using artificial intelligence and machine learning techniques. It is part of the project, "Developing Applications to Correct Jordanian Spoken Arabic to Proper Language Using Machine Learning Techniques." The paper addresses two interconnected dimensions. The computational linguistic dimension provides a technological framework for processing language using deep machine learning techniques. The linguistic dimension examines the interactions between colloquial Jordanian Arabic and Modern Standard Arabic (MSA). It analyzes changes in phonetics, morphology, orthography, syntax, rhetoric, and semantics. The study focuses on processing real-world texts derived from society's linguistic reality. It uses the Jordanian dialect as a comprehensive model for bridging colloquial and MSA. The goal is to develop applications that preserve and promote the use of eloquent and correct linguistic content. This applies to daily communication and written dialogues on technological platforms. The success of these applications in addressing the Jordanian dialect could inspire similar projects across other Arab countries. This would enhance cultural communication and linguistic preservation on a broader scale.

Keywords: Arabic language, natural languages processing, translation, Jordanian colloquial dialect, machine learning.