نحو تأسيس اللسانيات الحدّية الحاسوبية: رؤية رياضية–لسانية جديدة لتمثيل اللغة الطبيعية باستخدام المعادلات الجبرية الحدية

Towards a Foundational Framework for Computational Polynomial Linguistics
A Mathematical–Linguistic Vision for Representing Natural Language through Algebraic Polynomial Equations

د. جعفر يايوش *

جامعة عبد الحميد بن باديس–مستغانم

djaafar.yayouche@univ-mosta.dz

الاسم الكامل للباحث الثاني

مؤسسة الانتماء،

مخبر الانتماء

Email

تاريخ الاستلام: ٢٠٢٥/٠٧/٠٦

تاريخ القبول: ٢٠٢٥/٠٩/١٠.

النشر: ٢٠٢٥/١٢/١١

ملخص

تشهد اللسانيات الحاسوبية تطورًا متسارعًا بفعل التقدّم في الذكاء الاصطناعي والتعلّم الآلي، غير أنّ النماذج السائدة غالبًا ما تفتقر إلى تمثيل رمزي بنيوي قابل للتفسير، خاصة فيما يتعلّق بالبنى الصرفية والجذرية للغات الطبيعية كالعربية. تهدف هذه الورقة إلى اقتراح إطار تأسيسي جديد تحت مسمّى "اللسانيات الحدّية الحاسوبية"، يقوم على استخدام المعادلات الجبرية الحدودية (Polynomial Equations) لتحليل التحوّلات الصوتية والصرفية والبنيوية في النصوص. تقدّم هذه المقاربة منظورًا رياضيًا–لسانيًا يسمح بتمثيل العلاقات اللغوية عبر نماذج قابلة للحوسبة والحلّ والاشتقاق، ما يحيي منهج المعالجة الرمزية الدقيقة ضمن بيئة حاسوبية مرنة وفعالة. وتستعرض الورقة الأسس النظرية والرياضية لهذا الاتجاه، وتؤسس له بوصفه مسارًا معرفيًا واعدًا ضمن مشهد اللسانيات المعاصرة..

الكلمات المفتاحية: لسانيات حاسوبية، لسانيات حدّية، جبر حدّي، معادلات جبرية، تمثيل رمزي، بنية صرفية.

Abstract :

Computational linguistics has rapidly evolved under the influence of artificial intelligence and machine learning. However, most prevailing models lack symbolic and structurally interpretable representations, particularly in capturing morphological and root-based features of natural languages such as Arabic. This paper proposes a foundational framework termed Computational Polynomial Linguistics (CPL), which employs polynomial equations to model phonological, morphological, and structural transformations in language. The proposed approach introduces a novel mathematical–linguistic perspective that enables linguistic relations to be represented through computable, solvable, and derivable

algebraic models. The paper outlines the theoretical and mathematical foundations of CPL, positioning it as a promising paradigm in modern computational and theoretical linguistics.

KeyWords
Computational linguistics, polynomial linguistics, bounded algebra, symbolic representation, morphological structure.

*المؤلف المرسل

المقدمة:

١,١ الخلفية العامة

تطورت اللسانيات الحاسوبية منذ منتصف القرن العشرين عبر ثلاث مراحل رئيسية: النماذج الرمزية، ثم الإحصائية، وأخيرًا العصبية، مما أفضى إلى تطبيقات فعالة كتحليل النصوص والترجمة الآلية (Jurafsky & Martin, 2023, p. 3). إلا أن هذه النماذج، رغم كفاءتها، تفتقر إلى القدرة التفسيرية البنيوية الرمزية (Mohri et al., 2022, p. 145).

١,٢ السياق المعرفي العالمي

شهد تمثيل اللغة الطبيعية أربعة تيارات كبرى:

الاتجاه الإحصائي: استخدم نماذج مثل HMM وMaximum Entropy لمعالجة اللغة كاحتمالات متغيرة، لكنه أهمل البنية العميقة (Jurafsky & Martin, 2023, p. 55).

الاتجاه العصبي: اعتمد على الشبكات العصبية كـ BERT وGPT التي مثّلت اللغة كمتجهات عددية بلا بنية رمزية (Bender & Koller, 2020, p. 5186).

الاتجاه الرمزي: ركّز على النماذج القواعدية كـ CFGs، فوفّر وضوحًا بنيويًا، لكنه عانى من ضعف التوسّع، خاصة مع العربية (Pullum & Scholz, 2001, p. 33).

الاتجاه الجبري: استخدم الزمر والمونويدات لنمذجة اللغة رياضيًا، لكنه بقي نظريًا دون تطبيق فعلي (Berstel & Perrin, 1985, p. 12).

وفي ظل غياب توظيف المعادلات الحدّية الجبرية في التمثيل اللغوي، تقترح هذه الورقة تأسيس تيار خامس: اللسانيات الحدّية الحاسوبية (Computational Polynomial Linguistics)، لدمج الرمزية البنيوية والدقة الجبرية ضمن إطار قابل للتفسير والحوسبة.

1.3الأسباب والبواعث

رغم التقدم التقني الذي أحرزته نماذج التعلّم العميق في اللسانيات الحاسوبية، إلا أنه لم يُصاحب بنضج بنيوي يتيح تفسير الظواهر اللغوية رياضيًا أو رمزيًا .(Chomsky, 1965, p. 27) فالنماذج الإحصائية—العصبية تنجح في المهام الإجرائية لكنها تفتقر إلى نموذج تفسير داخلي يشرح نتائجها(Marcus, 2020, p. 6) ، ما أدى إلى أزمة "الصندوق الأسود" في الذكاء الاصطناعي اللغوي.

أما النماذج الرمزية، فتعاني من ضعف قابلية التوسيع(Pullum & Scholz, 2001, p. 147) ، إضافة إلى محدودية أدواتها في التمثيل الرياضي القابل للتنفيذ. وحتى النماذج الجبرية السابقة (كالزمر والمونويدات) ظلت نظرية ولم تُطبّق عمليًا في معالجة الظواهر الصرفية—الجذرية، رغم ما توفّره اللغة العربية من انتظام بنيوي يدعم هذا النوع من النمذجة .(Habash, 2010, p. 101)

من هنا تنشأ الحاجة إلى نموذج يتجاوز ثنائية "التمثيل الرمزي غير القابل للحوسبة" و"التعلم العددي غير المفسّر"، من خلال اعتماد المعادلات الحدّية الجبرية (Polynomial Equations) كأداة تمثيل تسمح بتحليل البنية والتحول، وتنفيذها حوسبيًا ضمن إطار رياضي—لساني متكامل.

كما يثير هذا التوجه سؤالًا إبستيمولوجيًا حول إمكانية تجاوز نموذج تورنغ التقليدي القائم على الثنائية الصفرية، نحو حوسبة رمزية—رياضية غير تقليدية تعتمد على حالات متعددة المعاني .(Multivalued Symbolic States) وهو ما يتقاطع مع أطروحات ما بعد تورنغ(Copeland, 2002; Siegelmann, 1999) ، التي تعيد النظر في مفهوم "الآلة" عندما تُدمج البنية الرياضية في آليات المعالجة.

1.4الإشكالية المركزية

في ظل محدودية النماذج العصبية في التفسير البنيوي، وانغلاق النماذج الرمزية، وعدم انتقال النماذج الجبرية التقليدية إلى التطبيق، تبرز إشكالية مركزية :**هل يمكن بناء نموذج جبري حاسوبي يمثل التحولات اللغوية البنيوية – صوتيًا وصرفيًا ودلاليًا – باستخدام معادلات قابلة للحل والتنفيذ؟**

السؤال يتحول إلى تحدٍ منهجي :**هل يمكن استخدام المعادلات الحدّية الجبرية كتمثيلات حوسبية رمزية–دلالية للغة؟** .لا يعني هذا فقط صياغة نموذج جديد، بل أيضًا إعادة تعريف المعالجة الحاسوبية للغة بوصفها نظامًا جبريًا متعدد الجذور، كما في نمذجة الصيغ الاشتقاقية العربية بمعادلات من الدرجة ***n***.

يشكّل هذا التوجه ما تُسميه الورقة بـ **"اللسانيات الحدّية الحاسوبية(CPL)** "، كتأسيس لحقل جديد يدمج الرمزية البنيوية بالدقة الجبرية ضمن إطار حوسبي قابل للتحليل.

إبستيمولوجيًا، تفتقر النماذج العصبية لقدرة *epistemic recoverability*، كما أشار(Marcus (2020, p. 5 ، حيث لا يمكن تتبع إنتاجها اللغوي، خلافًا لبنية اللغة التوليدية كما أبرزها(Chomsky (1965, p. 27 ، وهي بنية قابلة للتمثيل الرياضي.

استقرائيًا، تُظهر العربية انتظامًا صرفيًا ينبثق من الجذر الثلاثي، يُمكن تمثيله كمعادلة حدّية يعكس كل عامل فيها عملية صرفية، ما يُتيح التحليل، التنبؤ، والاستخراج العكسي.

رياضيًا، تعتمد النماذج الحالية على الإحصاء والمتجهات دون توظيف الجبر التحليلي الصريح. في المقابل، يسمح النموذج الحدّي بتمثيل العلاقات التحويلية بين الجذر والوزن بتحليل رياضي قابل للتنفيذ، ما يُعيد للحوسبة اللغوية بعدًا منطقيًا–تفسيريًا مفقودًا.

هذا السؤال الجوهري – **هل يمكن نمذجة اللغة كمعادلات حدّية قابلة للحوسبة؟** –لا يعيد فقط تعريف نمذجة اللغة، بل يطرح لغة رياضية تمثل البنية الاشتقاقية، متجاوزةً الانقسام التقليدي بين الرمزي والعصبي.

الفكر الرياضي العربي: من الخوارزمي إلى ابن البناء – اللغة كنَسَق رياضي حاسوبي

تستند اللسانيات الحدّية الحاسوبية إلى جذور معرفية راسخة في الفكر الرياضي العربي–الإسلامي، كما يتجلى في أعمال الخوارزمي وابن البناء المراكشي. فقد قدّم الخوارزمي في *الكتاب المختصر في حساب الجبر والمقابلة* تصورًا مبكرًا للجبر كأداة لحل المشكلات عبر معادلات قابلة للتحليل، مما مهّد لمفهوم "الخوارزمية" (الخوارزمي، ١٩٨٦؛ راشد، ١٩٩٨). وتُستعاد هذه الرؤية اليوم في نموذج يحوّل العلاقات الصرفية–الجذرية إلى معادلات قابلة للحوسبة.

أما ابن البناء، فقد وسّع هذا التصور في القرن الرابع عشر، مبرزًا العلاقات الرياضية كأنظمة رمزية قابلة للتجريد (ابن البناء، ١٩٨٦)، وهي بداية مبكرة لفكرة "التمثيل الرمزي". كما تُظهر علوم العَروض والنحو والبلاغة في التراث العربي تكاملًا بين البنية اللغوية والمنطق الرياضي، مما يعكس تصورًا للغة كنظام توليدي منظّم – وهو ما تسعى إليه النمذجة الحدّية الجبرية.

من هنا، لا تُعد المعادلات الجبرية الحدّية استعارة مستوردة، بل استعادة لرؤية إبستيمولوجية أصيلة تُعامِل اللغة كبنية رياضية قابلة للتحليل والتفسير.

زاوية التأسيس التأصيلي:

يُؤصّل هذا الامتداد المشروع الراهن بوصفه ليس فقط مقاربة حوسبية–جبرية تأسيسية، بل أيضًا تحديثًا للمنهج الجبري العربي ضمن الحوسبة الرمزية المعاصرة، ما يمنحه بعدًا معرفيًا مزدوجًا:

- **تأسيسيًا** في بنية النمذجة
- و**تأصيليًا** في مرجعيته الحضارية والمعرفية.

(1.4)الإشكالية المركزية

في ظل القصور البنيوي للنماذج العصبية، وانغلاق الرمزية، وعجز النماذج الجبرية الكلاسيكية عن التطبيق، تبرز الحاجة إلى نموذج رياضي حاسوبي جديد يمكّن من تمثيل الظواهر الصوتية والصرفية والدلالية عبر بنى قابلة للحل والتنفيذ. السؤال المطروح: هل يمكن استخدام **المعادلات الحدّية الجبرية** لتمثيل اللغة كنظام رياضي–رمزي قابل للحوسبة؟ هذه الإشكالية تتجاوز مجرد تمثيل لغوي، لتعيد تعريف المعالجة الحاسوبية نفسها، وتقترح تصورًا للغة كبنية **متعددة الجذور** تحاكي طبيعة المعادلات الجبرية. كما تنفتح على تساؤل تطبيقي مباشر: هل يمكن تحويل الصرف العربي إلى نموذج حدّي جبري قابل للتحليل البرمجي، بما يتفوق على قيود الأنظمة الحالية؟

تسعى الورقة للإجابة عبر تأسيس ما تسميه **اللسانيات الحدّية الحاسوبية(CPL)** ، كنموذج يتجاوز ثنائية الأداء والتمثيل، ويقدّم **معالجة لغوية قابلة للتفسير**.(**explainable computation**)

إبستيمولوجيًا، تظهر النماذج العصبية اعتمادًا على معرفة ضمنية غير قابلة للتفسير(Marcus, 2020, p. 5) ، خلافًا لطبيعة اللغة التوليدية القابلة للتحليل. (Chomsky, 1965, p. 27)

استقرائيًا، تكشف الصيغ العربية (فَعَلَ، فاعَلَ...) انتظامًا صرفيًا يمكن تمثيله كمعادلة من الدرجةn ، تمكّن من **التحليل، التنبؤ، والاستخراج العكسي**. **رياضيًا**، بخلاف النماذج التي تُعالج اللغة كنقاط في فضاء بيانات، يتيح **الجبر الحدّي** توصيف العلاقات التحويلية form - pattern - root عبر اشتقاق معادلات قابلة للتحليل والاستدلال.

بذلك، تطرح الورقة سؤالًا تأسيسيًا:

هل يمكن إنشاء نموذج **تمثيلي جبري** للعلاقات الصوتية–الصرفية–الدلالية، يُحوّل اللغة إلى نظام رياضي قابل للحوسبة؟

إن الإجابة تُعيدنا إلى فلسفة ما بعد–تورِنغ (Post-Turing Models) في **الحوسبة غير التقليدية** (Copeland,2000)؛(Siegelmann, 1999) ، حيث الحوسبة ليست ترميزًا ثنائيًا فحسب، بل **تمثيل رمزي–تفسيري** متعدد القيم. وهنا تبرز المعادلة الجبرية الحدّية كوسيلة لفهم اللغة، لا فقط لمعالجتها.

وفي هذا الإطار، تقترح الورقة مرحلة جديدة في تاريخ التمثيل الحاسوبي: بعد القواعد الصورية، والنماذج الإحصائية، والنماذج العصبية، تأتي **مرحلةCPL** ، حيث تُحوّل الظواهر اللغوية إلى معادلات قابلة للفك والتحليل، مدمجة بين الرمزية البنيوية والحوسبة التفسيرية.(Chomsky, 1965; Marcus, 2020)

بهذا، تُمثّل "اللسانيات الحدّية الحاسوبية" تحوّلًا جذريًا:من المحاكاة إلى الفهم، ومن الإحصاء إلى التحليل، ومن قواعد الجملة إلى **المعادلة البنيوية**، ومن سؤال "كيف تعمل النماذج؟" إلى "لماذا تعمل؟" و"كيف نفهمها؟"

(1.5)أهمية الموضوع: رياضيًا، إبستيمولوجيًا، حاسوبيًا، لسانيًا، وفلسفيًا

لا تكمن أهمية المشروع في تقديم بديل تقني للنماذج القائمة، بل في اقتراح نقلة مفاهيمية تعيد تعريف العلاقة بين **الرياضيات واللسانيات** .فبدل اختزال التحليل اللغوي في الإجراءات الإحصائية، يقترح المشروع مقاربة رياضية تفسيرية تُمكّن من فهم البنى اللغوية بمنطق جبري قابل للحوسبة.(Marcus & Davis, 2019, p. 196)

رياضيًا: يقترح المشروع استخدام **المعادلات الحدّية الجبرية** – (Polynomial Equations)كركيزة من ركائز الجبر الكلاسيكي – (Stewart, 2015) لتحليل التحولات الصرفية والاشتقاقية، بما يفتح أفقًا لتحويل اللغة إلى **نسق رياضي قابل للتوليد والحل** .(Grohe, 2021, p. 6)

إبستيمولوجيًا: أزمة التفسير في نماذج مثل GPT و (Bender & Koller, 2020) BERTتؤكد الحاجة إلى نموذج يُمكّن من **استرجاع بنية التحليل اللغوي وتفسيرها جبريًا،** مما يسدّ الفجوة بين القدرة على التوليد والفهم.

حاسوبيًا–تطبيقيًا: يُقدّم المشروع بديلًا تحليليًا للتمثيل الإحصائي التقليدي، باقتراح **فضاء جبري متعدد الجذور** يُمكّن من برمجة خوارزميات لغوية قائمة على التحليل الرياضي لا على تكرار الأنماط (Manning & Schütze, 1999; Jurafsky & Martin, 2023).

لسانيًا: يسمح **النظام الصرفي الجذري–الوزني** في العربية بتمثيل التحولات كمعادلات حدّية تعكس نمطًا اشتقاقيًا منتظمًا(Habash, 2010) ، ما يُعزّز دقة التحليل ويُقلل من اللبس والتجزئة.

رمزيًا–فلسفيًا: يستعيد المشروع رؤية **الخوارزمي وابن البناء** في اعتبار اللغة بنية رياضية توليدية (الخوارزمي، ١٩٨٦؛ ابن البناء، ١٩٧٤)، مقترحًا مقاربة تأصيلية تُحوّل اللغة من corpus إلى **معادلة دلالية–صرفية قابلة للحل،** تجمع بين الفهم الرمزي والمعالجة الرياضية.

(١,٦) أهداف الورقة البحثية

لا تقتصر هذه الورقة على معالجة تقنية جزئية، بل تهدف إلى إرساء منظور تأسيسي يُعيد ربط التمثيل اللغوي بالنموذج الجبري الحاسوبي، وذلك من خلال الأهداف التالية:

أولًا: التأسيس النظري لحقل جديد

تبلور مفهوم اللسانيات الحدّية الحاسوبية (Computational Polynomial Linguistics) كحقل يدمج بين البنية اللغوية والجبر التحليلي، متجاوزًا النماذج الصندوقية نحو تحليل قابل للتفسير (Grohe, 2021).

ثانيًا: اختبار التمثيل الجبري للصرف العربي

تحليل النماذج الصرفية العربية باستخدام معادلات متعددة الحدود، لتبيان قابلية الاشتقاق الجذري–الوزني للتمثيل الرياضي (Habash, 2010؛ Bender & Koller, 2020).

ثالثًا: نقد النماذج السائدة

تفكيك قصور النماذج العصبية من زاوية التفسير البنيوي، واقتراح بديل رياضي–رمزي يكشف البنية الداخلية عبر معادلات حدّية (Marcus & Davis, 2019).

رابعًا: بناء نموذج رياضي–لساني متعدد الأبعاد

صياغة نموذج يربط بين الجذر، الوزن، التحول الصوتي، والمعنى، داخل فضاء جبري تمثيلي غير خطي (Stewart, 2015).

خامسًا: تحفيز تطوير أدوات حوسبية جديدة

دفع عجلة بناء أنظمة تحليل صرفي وصوتي تعتمد على الحلول الجبرية، ما يُعزز التفسير ويُقلّل اللبس في معالجة النصوص (Manning & Schütze, 1999؛ Jurafsky & Martin, 2023).

(١,٧) المساهمة الأصلية للورقة البحثية: نحو اللسانيات الحدّية الحاسوبية (CPL)

لا تُعالج هذه الورقة تحسين النماذج القائمة، بل تقترح تأسيس إطار معرفي جديد تحت مسمى اللسانيات الحدّية الحاسوبية (Computational Polynomial Linguistics)، مبني على فرضية أن التحولات اللغوية قابلة للتمثيل كنظام جبري حدّي متعدد الجذور، قابل للحل الحاسوبي الرمزي.

تنقسم المساهمة إلى خمس مستويات:

١. مفاهيميًا

تضع الورقة أول تعريف صريح لـ CPL كحقل يُعيد توصيف الظواهر اللسانية عبر معادلات حدّية قابلة للتحليل، متجاوزًا النماذج السياقية أو العصبية، ومقتربًا من الشفافية الرياضية البنيوية (Grohe, 2021؛ Bender & Koller, 2020).

٢. إبستيمولوجيًا

تستجيب لأزمة الصندوق الأسود في NLP، بتقديم نماذج قابلة للاسترجاع التفسيري، حيث يمكن تتبع التحولات اللغوية ضمن بنى معادلات حدّية قابلة للحل (Marcus & Davis, 2019).

٣. رياضيًا–تمثيليًا

تقترح تحويل العلاقات الصرفية–الجذرية (كَفَعَلَ ← تَفَعَّلَ) إلى معادلات متعددة الحدود، تعبر كل مرحلة عن عامل صرفي–تحويلي، ما يفتح بابًا جديدًا لنمذجة عربية دقيقة (Habash, 2010؛ Beesley & Karttunen, 2003).

٤. تطبيقيًا

تُصاغ نماذج أولية لتحليل الصيغ الصرفية العربية باستخدام الجبر الحدّي، تشمل التوليد، الاسترجاع، والتنبؤ بالعلاقات، ما يُقدّم بديلاً للطرائق الإحصائية التقليدية.

٥. تاريخيًا–فلسفيًا

يُستأنف الإرث الجبري العربي (الخوارزمي، ابن البناء) بوصفه تصورًا تحويليًا للغة، لا فقط عدديًا، بل كنظام قابل للمعادلة الرمزية، مما يمنح المشروع تأصيلاً حضاريًا ومعرفيًا متماسكًا.

(١,٨) الإطار المنهجي

Foundational–Computational Synthesis

تعتمد الورقة منهجًا تركيبيًا يجمع بين التحليل النظري والتمثيل الحاسوبي، لصياغة **نموذج جبري–لساني جديد** يُعرف بـ **اللسانيات الحدّية الحاسوبية (CPL)**. ويهدف إلى تحويل البنية الصرفية–الصوتية للغة إلى نظام قابل للتحليل الرياضي والتنفيذ الرمزي.

تتوزع المنهجية على أربعة محاور:

إبستيمولوجي–رياضي

تُصاغ الظواهر الاشتقاقية داخل بنى جبرية متعددة الجذور، باستخدام أدوات من الجبر التحليلي ونظرية المعادلات الحدّية، لتوصيف التحولات اللغوية تمثيليًا وتحليليًا (Grohe, 2021; Stewart, 2015).

استقرائي–تحويلي

تُستقرأ الصيغ الصرفية من الجذور العربية وتُعاد صياغتها كدوال حدّية من الرتبة n، تُعالج كتحولات منتظمة قابلة للحوسبة (Ryding, 2005; Habash, 2010).

تركيبي–حاسوبي

تُطوَّر نماذج أولية باستخدام أدوات الجبر الرمزي (مثل SymPy)، تحوّل المعادلات إلى خوارزميات اشتقاقية قادرة على الاسترجاع والتفسير البنيوي (Beesley & Karttunen, 2003).

مقارن–نظري

تُقارن CPL بالنماذج العصبية والرمزية، من حيث البنية، التفسير، القابلية الحاسوبية، والشفافية الإبستيمولوجية، مُبرزة تفوق CPL من حيث التكامل بين الدقة الرمزية والمرونة التنفيذية (Marcus & Davis, 2019).

يُعيد هذا الإطار تعريف تمثيل اللغة عبر ثلاث مستويات: تحليل رياضي دقيق، تنفيذ حوسبي تجريبي، وتأصيل مفاهيمي مستقل، ما يجعل CPL نموذجًا قابلًا للتوسيع نحو لغات أخرى، وأداة لتطوير معالجات لغوية مبنية على المنطق الجبري القابل للتفسير.

(١,٩) الفروض العلمية والنموذج النظري

أولًا: الفروض العلمية

ترتكز الورقة على ثلاث فرضيات قابلة للاختبار:

القابلية الجبرية للتمثيل اللغوي

تفترض إمكانية نمذجة الظواهر الصرفية والدلالية بمعادلات حدّية من الرتبة **n**، استنادًا إلى الجبر الكلاسيكي (Stewart, 2004) والنحو الوزني العربي (Wright, 1996؛ ابن جني)..

تعويض الجبر الرمزي عن الشبكات الإحصائية

تفترض قدرة النموذج الجبري على تقديم تمثيل لغوي شفاف بديلًا للنماذج العصبية غير التفسيرية (Bender & Koller, 2020؛ Chomsky, 1957).

الحوسبة الرمزية للاشتقاق

يفترض أن CPL قادر على محاكاة الاشتقاقات الصرفية رياضيًا دون الاعتماد على البيانات الضخمة، مستندًا إلى النمذجة الوزنية التراثية (Habash, 2010؛ الخليل، ابن جني).

ثانيًا: النموذج النظري الرياضي ثانيًا: النموذج النظري الرياضي
Algebraic Theoretical Framework

يقوم النموذج النظري على توصيف العلاقات الاشتقاقية والتحويلية في اللغة كأنظمة متعددة الجذور، تُعبّر عنها عبر معادلات حدّية تأخذ الشكل التالي:

$$K = (d + c)(x + b)(x + a)(x + x) = P(x)$$

حيث:

- x المتغير البنيوي الذي يمثل الجذر الصرفي الأولي
- a ، b، c، d: عوامل التحول الصوتي أو الإضافي (مثل: الإعلال، الزيادة، النقل، الإدغام)
- K **الناتج السطحي النهائي(surface form) ، والذي قد يُعبر عن صيغة صرفية أو دلالية مركبة.**

ثالثًا: مثال على النمذجة الجبرية للاشتقاق العربي

لنأخذ الجذر الثلاثي ك–ت–ب كمثال:

- **الصيغ :كتب، كاتب، مكتوب، استكتب**

- **يُمكن تمثيل هذه الصيغ كحلول لمعادلة من الشكل:**

$$y = (\delta + \gamma)(x + \beta)(x + \alpha)(x + x) = P(x)$$

بحيث:

X : يمثل البنية الجذرية "كتب"

α,β... تحولات صرفية (إدخال همزة، زيادة تاء، نقل بنية، إلخ)

Y : البنية السطحية (مثل: استكتب)

هذا يسمح بتحليل الاتجاهين:

من الجذر إلى الوزن (اشتقاق)

من الصيغة إلى الجذر (تحليل عكسي)

رابعًا: خصائص النموذج المقترح

Model Properties

١. **الاشتقاق العكسي** :(**Reverse Derivation**) القدرة على استرجاع البنية الأصلية من الصيغة النهائية عبر تحليل المعادلة.

٢. **الامتداد متعدد اللغات** :(**Multilingual Portability**) قابلية التعميم على لغات ذات نظام اشتقاقي شبيه، مثل العبرية، الأمهرية، والأكادية.

٣. **الكفاءة في بيئات منخفضة الموارد** :(**Low-Resource Compatibility**) يُمكن تشغيل النموذج في غياب البيانات الضخمة، لارتكازه على القواعد البنيوية بدل التعلم من التكرار.

٤. **تكامل رمزي–بنيوي** :(**Symbolic–Structural Integration**) يجمع بين التحليل البنيوي الدقيق والتمثيل الرمزي القابل للتفسير، ما يُعيد التوازن بين الفعالية والتفسير في NLP.

القسم ٢: الإطار النظري والمفاهيمي (Theoretical and Conceptual Framework):

2.1 الجبر في الحوسبة اللغوية: الخلفية النظرية والتاريخية

(Algebra in Computational Linguistics: A Historical–Theoretical Overview)

(أ) المدخل النظري: إعادة الاعتبار للجبر كنموذج بنيوي–تفسيري

رغم أن الجبر شكّل عبر التاريخ أحد الأعمدة المؤسسة للتمثيل البنيوي في العلوم الصورية، إلا أن حضوره ضمن علم اللغة الحاسوبي ظل هامشيًا، غالبًا داخل أطر محصورة بالنماذج الرمزية الصورية (Symbolic–Formal

Approaches). بعد هيمنة استمرت لعقود للنماذج الإحصائية ثم العصبية، عادت في السنوات الأخيرة أطروحات الحوسبة الرمزية والجبرية إلى الواجهة، بدافع الحاجة إلى نماذج قابلة للتفسير البنيوي والاشتقاق الرياضي، بدلًا من الاقتصار على الأداء التنبؤي (Menzel, 2010; Marcus & Davis, 2019).

يمثل الجبر – لا سيما بصيغه التحليلية والحدّية – أداة قوية لإعادة صياغة الظواهر اللغوية ضمن نماذج قابلة للحل، توليديًا وتحويليًا، في آنٍ معًا.

(ب) الجذور المعرفية للعلاقة بين الجبر واللسانيات

بدأ الربط بين الجبر واللغة مع النحو التوزيعي لهاريس، وتبلور مع النحو التحويلي عند تشومسكي (١٩٥٧) الذي قدّم اللغة كنظام توليدي قابل للتمثيل الرياضي.

وفي السبعينيات، ظهرت اللسانيات الجبرية (Algebraic Linguistics) التي استخدمت المجموعات والزمر والحلقات لتحليل البنى الصرفية والنحوية (Grätzer, 2008)، لكن دون تطبيق فعلي في حوسبة اللغة الطبيعية.

(ج) الحوسبة الجبرية الرمزية

يندرج الجبر الرمزي ضمن أنظمة الجبر الحاسوبي (Computer Algebra Systems) المستخدمة في مجالات علمية كبرى، مثل Mathematica وSymPy. ومع أنه استُخدم في تمثيل المعاني وتركيب العبارات (Pollard & Sag, 1994)، فقد ظل بعيدًا عن مستويات الحوسبة الاشتقاقية أو الصرفية، مما أظهر فجوة بين النظرية الجبرية والتطبيق اللغوي.

(د) التحديات الإبستيمولوجية

تُفسر محدودية استخدام الجبر في اللسانيات باعتقاد سائد أن اللغة نظام مفتوح وسياقي لا يمكن احتواؤه رياضيًا (Hale & Reiss, 2008). لكن هذا يُغفل إمكانية توصيف المستويات البنيوية – خصوصًا الصرف والاشتقاق – بدقة رياضية، كما أوضحتها أعمال Chomsky (١٩٥٧) وBeesley & Karttunen (٢٠٠٣).

(هـ) الجبر كنموذج تفسيري

في مواجهة نماذج الصندوق الأسود، ظهرت دعوات لتبنّي نماذج تفسيرية رمزية–جبرية (Kornai, 2007; Pollard & Sag, 1994). وتقترح الورقة الحالية استخدام المعادلات الجبرية الحدّية (Polynomial Algebra) لتحليل وتوليد البنى الصرفية–الاشتقاقية، مما يفتح أفقًا جديدًا في علم اللغة الحاسوبي الرمزي.

٢,٢ اللسانيات الجبرية – الملامح النظرية والأسس الإبستيمولوجية

(أ) تمهيد مفاهيمي

تُعيد اللسانيات الجبرية تأطير اللغة كبنية رياضية، بديلة للنماذج الشكلانية والإحصائية، بالاعتماد على فرضية أن الظواهر اللسانية يمكن تمثيلها عبر بنى جبرية مثل الزمر، الحقول، والمعادلات متعددة الحدود (Kornai, 2007؛ Marcolli et al., 2023).

(ب) البنية المفاهيمية

التمثيل الرمزي: تُعامل القواعد النحوية كعلاقات تحويلية قابلة للتحليل داخل هياكل جبرية (Partee et al., 1990).

الجبر التوليدي: تُطبّق معاملات صرفية–دلالية باستخدام (Pollard & Sag, 1994) morphisms.

التصنيف الظاهراتي: تصنَّف الظواهر اللسانية داخل طبقات رياضية مثل الزمر الجزئية (Grätzer, 2008).

(ج) من النماذج الكلاسيكية إلى البنية الحدّية

انتقل الحقل من النماذج التوزيعية إلى النحو الصوري (CFG, HPSG)، دون شفافية تفسيرية كافية (Shieber, 1985). أما النماذج الحديثة (Merge ،Hopf Algebras) فتمثّل اتجاهًا نحو تأويل تركيبي–رياضي (Marcolli et al., 2023). يُقترح هنا نموذج حدّي متعدد الجذور يُمثّل كل اشتقاق كحل لمعادلة قابلة للتفسير والتحليل.

جدول مقارنة:

مجالات التطبيق	الخصائص الجبرية	النموذج
الاشتقاق وتمثيل السياق	متجهات/مصفوفات	النموذج التوزيعي
توليد الجمل	تحويل جبري رمزي	CFG, HPSG
البنية النحوية	مخططات تركيبية	التحليل الجبري
الترجمة والبحث	طبقات جبرية صورية	اللغويات الرياضية

(د) المنهج الاستقرائي البنيوي: يعتمد النموذج الحدّي على استقراء الانتظامات الصرفية–الصوتية مثل "تَفَعَّلَ" و"مُفَاعِل"، وتحليل العلاقة بين الجذر والتحول كمعامل جبري ضمن معادلة حدّية، مع توصيف الشكل السطحي كنتيجة رياضية قابلة للتحقق الرمزي والاسترجاع البنيوي.

(هـ) فرضيات مركزية في اللسانيات الجبرية التأسيسية

ترتكز المقاربة الجبرية على فرضيات قابلة للاختبار:

- إمكانية نمذجة الظواهر اللغوية رياضيًا عند مستوى مناسب من التجريد
- الاشتقاق الصرفي والتحول الدلالي نواتج لعوامل جبريّة ضمن معادلات حدّية
- توليد وتحليل البنى اللغوية ممكن دون اعتماد على النماذج الاحتمالية أو العصبية
- قابلية تعميم النموذج الحدّي إلى مستويات تركيبية أعلى

(و) العلاقة بالمشروع الحالي: من الجبر الخطي إلى الجبر الحدّي الرمزي

يقترح المشروع الحالي تجاوز البنى الخطية (مثل المتجهات) إلى نموذج جبري حدّي متعدد الجذور، حيث:

- يُمثَّل الجذر كمتغير أولي
- تُعامل التحولات الصرفية كعوامل جبريّة
- يُمكن توليد الأشكال الصرفية وتحليلها عكسيًا
- يُدمج التفسير الرمزي داخل نموذج قابل للتنفيذ
- يُطبّق على لغات اشتقاقية مثل العربية، التركية، واللغات السامية

تشومسكي والمنعطف الجبري المعاصر

انطلق تشومسكي في Syntactic Structures (١٩٥٧) من تأطير اللغة كنظام صوري توليدي يعتمد على قواعد إنتاجية وتحليل قائم على نظرية اللغات الشكلية، لكنه ظل مقيدًا ببنى تتابعية.

لاحقًا، ومع الحاجة إلى نماذج تفسيرية تتجاوز قصور النماذج الإحصائية والشبكية (Bender & Koller, 2020)، اقترح تشومسكي مع Berwick وMarcolli نموذج البنية الجبرية لعملية الدمج (Algebraic Structure of Merge)، حيث تُعاد صياغة Merge كعملية جبرية ضمن هياكل زمرية أو تصنيفية (Marcolli et al., 2023).

يمثل هذا التحول انتقالًا من تسلسل القواعد إلى تحويلات قابلة للحل داخل منظومات رياضية تفسيرية، تفتح أفقًا لبناء نماذج لغوية قابلة للحوسبة والتعميم الرمزي.

ضمن هذا السياق، يأتي مشروع اللسانيات الحدّية الحاسوبية كتطوير لهذا التحول، مقترحًا بنى متعددة الجذور ومتغيرة المتغيرات (Multivariate Polynomial Systems) لتمثيل التحولات الصرفية والدلالية داخل نموذج جبري يمكن تفسيره وتحليله.

هذا لا يُكمل مسار تشومسكي فحسب، بل يوسّعه نحو نماذج لغوية قائمة على المعادلة والتحويل الجبري بدلًا من المحاكاة أو التسلسل القاعدي فقط.

٣. مقارنة نقدية مع النماذج الإحصائية والعصبية

رغم نجاح النماذج الإحصائية والعصبية في مهام المعالجة الآلية، فإنها تفتقر إلى التمثيل البنيوي العميق والتفسير الرياضي الصريح للتحولات اللغوية (Bender & Koller, 2020; Marcus, 2022).

النماذج الإحصائية

تعتمد على سلاسل Markov وN-grams، لكنها تفتقر إلى البنية الصورية وتعالج اللغة عبر تكرار الأنماط (Jurafsky & Martin, 2023). وقد رفض تشومسكي هذه المقاربة لعدم قدرتها على تفسير الكفاءة اللغوية (Chomsky, 1957).

النماذج العصبية

مثل GPT وBERT تعتمد على embeddings دون تفسير داخلي لآليات التوليد، وتبقى صناديق سوداء بلا شفافية بنيوية (Belinkov & Glass, 2019; Marcus, 2022).

النموذج الجبري–الحدّي المقترح

يُقدّم بديلًا صوريًا قائمًا على معادلات متعددة الجذور، يُمكّن من التحليل والتوليد والتفسير الرمزي (Grohe, 2021; Marcolli et al., 2023).

المعيار	النماذج الإحصائية / العصبية	النموذج الجبري–الحدّي
البنية	غائبة أو ضمنية	صريحة وقابلة للتحليل الرمزي
التمثيل الدلالي	سياقي	علاقات رياضية داخلية
قابلية التفسير	محدودة(Black-Box)	عالية(Symbolic Traceability)
آلية التوليد	تعتمد على بيانات ضخمة	تحويلات قابلة للحل
البنية النظرية	غير صورية أو غير تفسيرية	صورية – رياضية – تجريدية

الفرضية التأسيسية: اللغة ليست احتمالات أو شبكات فقط، بل نظام جبري توليدي يحتوي على قواعد داخلية قابلة للتمثيل والتحليل الرياضي الكامل.

٣,١ التحليل الإبستيمولوجي–التطبيقي المقارن

(أ) النماذج الإحصائية

تعالج اللغة كسلاسل خطية باستخدام N-gram وHMM، وتُقدّر احتمالات الانتقال بين الكلمات. رغم فعاليتها في الترجمة والإكمال، فهي لا تولّد بنية نحوية أو تمثل المعنى البنيوي، وتعتمد كليًا على بيانات التدريب (Chomsky, 1957).

(ب) النماذج العصبية

مثل GPT وBERT، تعتمد على embeddings ومعماريات attention، وتُنتج نتائج لغوية بدون تفسير داخلي أو تمثيل نحوي واضح، وتُعد صناديق سوداء غير قابلة للفهم أو البرهنة (Belinkov & Glass, 2019; Marcus, 2022).

(ج) النماذج الرمزية الشكلية

تشمل CFG وHPSG، وتُحلل البنية بدقة منطقية، لكنها تعاني من محدودية في التوليد، وتعميم التفسير على لغات مختلفة (Pollard & Sag, 1994).

(د) النموذج الجبري–الحدّي: تحليل رياضي وتوليد صوري قابل للتفسير

يمثل النموذج المقترح خطوة مفاهيمية جديدة: اعتبار اللغة نظامًا جبريًا تُشتق فيه البنى الصرفية والنحوية عبر معادلات حدّية متعددة الجذور. مثال رياضي توضيحي:

إذا كانت x تمثل الجذر:

$$c + bx + {}^{2}ax = f(x)$$

فإن $f(x)$ تُنتج الشكل الصرفي الجديد (اسم، فعل، صيغة صرفية).

هذا النموذج:

- يُمكّن من صياغة العمليات اللغوية في بنى قابلة للتحقق رياضيًا.
- يُوفّر تفسيرًا صريحًا لعلاقات المعنى–التركيب.
- يُعمّم على أنظمة صرفية–نحوية متعددة اللغات.

(هـ) مقارنة تحليلية

البُعد التحليلي	النماذج الإحصائية	النماذج العصبية	النماذج الرمزية	الجبري–الحدّي
التمثيل البنيوي الصريح	✗	✗	✓	✓✓
القدرة التوليدية الرياضية	✗	محدود	محدود	✓✓
تحويل الظواهر إلى معادلات	✗	✗	جزئي	✓✓
التفسير الرياضي–الصوري	✗	✗	✓	✓✓
تحليل دلالي داخلي	سطحي	متغير	جزئي	✓✓
تعميم على لغات متعددة	ضعيف	مقيد بالبيانات	محدود	✓✓
قابلية التحقق الحوسبي البرهاني	✗	✗	محدود	✓✓

(و) الخلاصة المفاهيمية

النموذج الجبري–الحدّي لا يُعيد فقط تعريف أداء النماذج، بل يُغيّر النظرة إلى اللغة ذاتها: ليست تكرارًا أو تعلمًا عصبيًا، بل نظام رياضي رمزي قابل للتحليل، البرهنة، والتجريد البنيوي.

2.4 الجبر الحدّي كنموذج إبستيمولوجي في النمذجة اللغوية (Polynomial Algebra as an Epistemological Framework for Symbolic Linguistic Modeling)

(أ) من الجبر كأداة إلى الجبر كنموذج معرفي

لم يعد الجبر الحدّي يُعد مجرد أداة لحل المعادلات، بل أصبح يُفهم كنموذج إبستيمولوجي يعيد توصيف الظواهر اللغوية ضمن بنى رياضية قابلة للتحليل الصوري، وفق(1997) Rota ، الذي يرى أن جوهر الرياضيات يكمن في تمثيل البنى المعرفية بصيغ رمزية قابلة للتحقق. من هنا يُقترح تصور اللغة كنظام جبري–تحويلي تتجسد فيه مكوناتها كدوال حدّية.

(ب) الأسس الرياضية: البنية والدالة

يتعامل الجبر الحدّي مع دوال من الشكل:

$$_0a + \ldots + {}^{n-1}a_{n-1}x + {}^{n}a_{n}x = P(x)$$

ويُعاد توظيفها لتمثيل التتابعات اللغوية التحويلية كبُنى رياضية–رمزية قابلة للتعميم الحوسبي.

(ج) في الاشتقاق والتحليل:

- **الاشتقاق الصرفي:** الجذر "كتب" يُشتق منه: "كاتب"، "كتابة"، إلخ، بدالة صرفية:

$$_na_nx + \ldots + _2a_2x + _1a_1x = S(x)$$

حيث كل a_i يمثل تحولًا صرفيًا.

التحليل النحوي:

العلاقات مثل "فاعل–فعل–مفعول" تُنَظَّم داخل دوال تُحاكي البنية الجملية عبر تحويلات على المتغيرات اللغوية.

الدلالة المركبة:

وفقًا لمبدأ Frege، تمثل العلاقات الدلالية بدوال متعددة المتغيرات:

$$a_3z + a_2y + a_1x = M(x, y, z)$$

(د) الجبر الخطي مقابل الحدّي:

الجبر الحدّي	الجبر الخطي	الجانب
بنى متعددة الرتبة	فضاء شعاعي	**البنية**
غير خطية، تحويلية	خطية	**العلاقات**
معقدة، تداخلية	بسيطة	**النمذجة**
ديناميكي	ثابت	**التوليد**
عالية	محدودة	**الملاءمة اللسانية**

يُبرز هذا التفوق قدرة الجبر الحدّي على توليد التراكيب اللغوية الرمزية وتمثيل التعقيد البنيوي.

(هـ) نحو نموذج لغوي متعدد المستويات

تمكن هذه النمذجة من ربط البنية والمعنى والسياق ضمن إطار رمزي–صوري موحَّد.

(و) الامتداد التاريخي العربي

يرتبط النموذج بجذور عربية أصيلة:

- **الخوارزمي** :التأسيس النظري للجبر.
- **ابن البناء** :التحويلات العددية كأنماط صرفية.
- **الكرجي** :معادلات متعددة الرتب قريبة من النموذج الحدّي المعاصر. وهكذا، يُستعاد الجبر الحدّي كتقليد إبستيمولوجي متجذّر في التراث العلمي الإسلامي، يعيد بناء العلاقة بين اللغة والمعنى عبر الرياضيات الرمزية.

(أ) من الوصف اللغوي إلى التجريد الرياضي

شهدت اللسانيات منذ منتصف القرن العشرين تحولًا إبستيمولوجيًا نحو النمذجة الشكلية–الرياضية، بفضل أعمال **Chomsky (1956)** التي أسّست للنحو التوليدي عبر مفاهيم **Finite Automata** و**Formal Grammars**. ورغم قدرتها على توصيف البنية السطحية، ظهرت محدوديات في تمثيل السياق والدلالة المركبة، مما مهّد لظهور نماذج جديدة تُعيد تمثيل اللغة كبنية تحويلية–رمزية قابلة للقياس الرياضي.

(ب) خرائط النماذج الرياضية

مساهمة الجبر الحدّي	القصور	نقطة القوة	النموذج
تمثيل التحولات كبُنى تحويلية من الرتبة العليا	ضعف دلالي–سياقي	توصيف صوري للتحولات الشكلية	Automata Theory
إعادة كتابة القواعد كمعادلات متعددة الحدود	غياب العمق الدلالي	قواعد تركيبية صارمة (Chomsky Hierarchy)	Formal Grammars
تفسير الأنماط الإحصائية داخل نسق جبري	عدم التفسير الرمزي	تعلم احتمالي(Markov, N-gram)	Statistical Models
نمذجة صرف–تركيب–دلالة كعلاقات متعددة الطبقات	ضعف بنيوي–دلالي	تمثيل متجهاتي(Word2Vec, GloVe)	Linear Algebra Models

(ج) النموذج الجبري الحدّي: تكامل رياضي–رمزي

يقترح النموذج الجبري الحدّي إعادة تأويل العلاقات الصرفية–التركيبية–الدلالية عبر معادلات رياضية متعددة المتغيرات، ما يسمح بفهم اللغة كبنية رياضية قابلة للقياس والتفسير:

$$L(x, y, z) = \alpha P(x) + \beta Q(y) + \gamma R(z)$$

حيث:

$P(x)$: دالة التحولات الصرفية للجذر.

$Q(y)$: دالة البنية التركيبية النحوية.

$R(z)$: دالة الدلالة والسياق.

α, β, γ: معاملات قابلة للضبط التداولي والخطابي.

هذا النموذج لا يُلغي النماذج السابقة، بل يُعيد دمجها في إطار بنيوي–رياضي أوسع يمكّن من الحوسبة الرمزية والتحقق الصوري للغة.

(د) نموذج تطبيقي تحليلي: تفكيك بنيوي–رمزي لجملة عربية

تمثل جملة" :**الطفلُ يركضُ في الحديقةِ** ".نموذجًا بسيطًا لتحليل لساني متعدد الطبقات، يمكن تمثيله وفق أربعة نماذج:

النموذج	التمثيل البنيوي
Automaton	تسلسل تحويلي محدود(Finite-State Transitions)
Formal Grammar	قواعد إنتاج من النوعContext-Free
Linear Algebra	تمثيل الكلمات كمتجهات باستخدام(Word2Vec (Mikolov et al., 2013
Polynomial Algebra	دمج صرف، تركيب، ودلالة ضمن معادلة متعددة المتغيرات

يسمح هذا التمثيل الأخير بإدماج **الصرف**، **التركيب**، **والدلالة** في معادلة واحدة، مستلهمًا الأسس الجبرية عند **الخوارزمي** و**ابن البنّاء** .وبهذا يُعاد تشكيل بنية الجملة كنسق رياضي–رمزي قابل للحوسبة، بما يعكس مفهوم "**اللغة كوحدة تحليلية متعددة الطبقات**" .(Jurafsky & Martin, 2009)

(هـ) نحو نموذج لساني متعدد الطبقات

النموذج الجبري الحدّي يتيح بناء **نظام لساني تكاملي** بخصائص:

١. **تكامل نحوي–دلالي** في دالة واحدة.

٢. **تحولات صرفية–نحوية** قابلة للتمثيل كمعادلات من الرتبة العليا.(Cox et al., 2015)
٣. **توليد وتحليل آلي** متعدد المستويات.(Jurafsky & Martin, 2009)
٤. **قابلية التوسيع اللغوي** إلى لغات اشتقاقية كالعبرية والسريانية عبر ضبط المعاملات فقط (Rashed, 1996).

(و) العربية كنموذج رياضي–اشتقاقي

تُعد العربية بيئة مثالية للنموذج الجبري الحدّي، بفضل نظام الجذر والوزن الذي يُمكّن من صياغة الاشتقاق ضمن دوال رياضية، بخلاف النماذج الإحصائية التي تعجز عن تمثيل هذا العمق البنيوي. ويمكن صياغة ذلك بالدالة التالية:

$$f_i(r, w) \cdot {}_i a \sum_{i=1}^{n} = F(w)$$

حيث:

- r: الجذر اللغوي
- w: الوزن الصرفي
- if: دوال التحويل الاشتقاقي
- ia: معاملات التمثيل الرمزي

تمثيل الجملة كدالة حدّية متعددة الحدود

(1)من البنية اللغوية إلى البنية الحدّية

تقترح الورقة تحويل الجملة إلى **دالة حدّية** يمثل كل حدّ فيها مكوّنًا لسانيًا (فاعل، فعل، ظرف...)، حيث يُعبر كل معامل عن وزن تركيبي أو دلالي معين، ضمن تصور أن البنية اللغوية متعددة الطبقات (Kornai, 2007; Marcus, 1967):

$$L(x, y, z) = \alpha P(x) + \beta Q(y) + \gamma R(z)$$

(٢) توزيع لغوي للدالة

$P(x)$: التحولات الصرفية
$Q(y)$: البنية النحوية
$R(z)$: الدلالة والسياق
α, β, γ: معاملات قابلة للضبط

(٣) التمثيل النمطي للجمل

الجملة	النموذج الحدّي
اسمية	حدّ بسيط
فعلية	حدّ من رتبة ثانية
مركبة	متعدد الحدود
شرطية أو متعدية المعنى	حدّي شرطي

(٤) الطبقات التوزيعية

يتم تمثيل الجملة عبر أربع طبقات:

النحو (Syntactic)، الدلالة (Semantic)، السياق (Pragmatic)، التوزيع (Statistical)

يتيح ذلك دمج (Mohri, 2009) weighted automata ونماذج (Pollard & Sag, 1994) HPSG ضمن التمثيل الحدّي.

(٥) الأساس النظري

يقف النموذج على أعمال:

Marcus (١٩٦٧): النحو الجبري، Kornai (٢٠٠٧): اللسانيات الرياضية، Pollard & Sag (1994): HPSG، Mohri (٢٠٠٩): الآلات المثقلة

٢,٧: الجذر والصيغة كنموذج جبري

(١) الجذر متغير رياضي: يعاد تأويل الجذر العربي كتعبير رياضي توليدي مستلهم من الخليل بن أحمد:

$$^{3}a_{3}x + {}^{2}a_{2}x + a_{1}x + {}_{0}a = W(x)$$

(٢) الصيغ كتحويلات على الجذر: كل صيغة = دالة تطبيقية على الجذر:

مثال: مكتوب=fمفعول(كتب)

(٣) توافق النموذج مع نظام الخليل

الخليل	التمثيل الجبري
الجذر	x
الوزن	f(x)
الاشتقاق	$f^n(x)$

(4)مكاسب النموذج

١. **تعريب** formal morphology
٢. **تكامل حوسبي** مع NLG وNLP
٣. **تمثيل متكامل** Morphology–Syntax–Semantics

(5)دراسات داعمة

- Beesley ، Marcus (1967) ، Bohas (1990)، Boudelaa & Marslen-Wilson (2001) & Karttunen (2003)

3: التطبيقات الحاسوبية للنموذج الحدّي

3.1: تمثيل الكلمة جبريًا

$$^3\delta x + {}^2\gamma x + \beta x + \alpha = F(x)$$

3.2 :نموذج أولي بلغة Python

تحويل الجذر إلى صيغة صرفية قابلة للتخصيص.

```python
def polynomial_root(x):
    return [x[0] + "a", x[1] + "u", x[2] + "i"]

def maktoub_form(x):
    root = polynomial_root(x)
    return "m" + root[0] + root[1] + "oo" + root[2]

print(maktoub_form(["k", "t", "b"]))  # "تقريب لطقي لـ "مكتوب
```

٣,٣: النموذج الطبقي للتوليد

الطبقة	التوصيف
1	الجذر xx
2	تحويل f(x)
3	دالة حدّية P(f(x))
4	إسقاط دلالي ∫P(x)dxf
5	الجملة Σ

٣,٥: نتائج تطبيقية:

الجذر	الكلمة	الوزن	التمثيل الجبري
ك–ت–ب	مكتوب	مفعول	f مفعول(كتب)
خ–ر–ج	خارج	فاعل	f فاعل(خرج)

الجذر	الكلمة	الوزن	التمثيل الجبري
ع–ل–م	استعلام	استفعال	fاستفعال(علم)

٣,٦: قابلية الإدماج في NLP:

SpaCy + وحدات صرفية مخصصة، OpenFST للتحليل الثنائي، TensorFlow (نماذج رمزية)

٤: تقييم النموذج الحدّي

٤,١: القوة التفسيرية: يمثل النموذج بديلًا شفافًا للنماذج الإحصائية (Black-box)، عبر دوال صرفية قابلة للبرهنة والتوليد.

٤,٢: القيود

رياضية: تعقيد المعادلات العليا، حوسبية: الحاجة لخوارزميات دقيقة، لسانية: محدودية في اللهجات غير النظامية

٤,٣: أسئلة مفتوحة: درجة التعقيد المثلى؟ تعميم على لغات غير اشتقاقية؟ مكتبة معيارية للجذور؟

٤,٤: آفاق التكامل: مع الشبكات العصبية: التعلم الرمزي–الجبري، مع النماذج الرمزية: توحيد التحليل والتوليد مع Formal Grammars: تحويل القواعد إلى معادلات.

٤,٥: تطبيقات مستقبلية: نظم تعليم صرفي، مكتبة جبرية مفتوحة، وحدات ترجمة آلية عربية، محرك صرفي مبني على المعادلات الحدّية.

٤,٦: خاتمة نظرية: يُعد النموذج الجبري الحدّي صيغة معاصرة لدمج الرؤية الفراهيدية في بنية رياضية قابلة للحوسبة. وهو لا ينافس النماذج التقليدية، بل يُكملها بطبقة تفسيرية دقيقة تعيد صياغة العلاقة بين الجذر والصيغة كنظام رمزي–عددي قابل للتنفيذ.

رغم أن الأدوات القائمة مثل Farasa وCAMeL تعتمد على قواعد أو تعلم آلي، إلا أنها تفتقر إلى تمثيل جبري–صرفي صريح. بالمقابل، يوفر هذا النموذج إطارًا يمكن دمجه معها لتفسير وتحليل البنية الصرفية بدقة قابلة للتحقق البرمجي، مما يمهّد لبناء أنظمة لغوية تفسيرية قابلة للتوسعة، خصوصًا في اللغات الاشتقاقية كاللغة العربية.

قائمة المراجع:

بالكامل وبمنهجية صارمة، إليك القائمة الموحدة للمراجع مرتبة حسب فئاتها وبـ تسلسل رقمي واحد شامل، وفق تنسيق APA 7 ، مع احترام الترتيب الأبجدي داخل كل فئة:

أولًا: الكتب العربية:

١. ابن البناء المراكشي، أحمد .(1974) .*العدد المجهول* .تحقيق: رشدي راشد. بيروت: دار الغرب الإسلامي.

٢. ابن البناء المراكشي، أحمد .(1986) .*تلخيص أعمال الحساب* .تحقيق: مصطفى ملا مسلم. بيروت: دار الغرب الإسلامي.

٣. الخوارزمي، محمد بن موسى .(1986) .*الكتاب المختصر في حساب الجبر والمقابلة* .تحقيق: رشدي راشد. حلب: معهد التراث العلمي العربي، جامعة حلب.

٤. الخوارزمي، محمد بن موسى .(1986) .*الكتاب المختصر في حساب الجبر والمقابلة* .تحقيق: فؤاد سزكين. فرانكفورت: معهد تاريخ العلوم العربية والإسلامية.

٥. الفراهيدي، الخليل بن أحمد. (د.ت) .*كتاب العين* .تحقيق: مهدي المخزومي وإبراهيم السامرائي. بيروت: دار ومكتبة الهلال

ثانيا: الكتب الأجنبية:

6. Beesley, K. R., & Karttunen, L. (2003). *Finite-State Morphology*. CSLI Publications.
7. Berwick, R., & Chomsky, N. (2016). *Why Only Us: Language and Evolution*. MIT Press.
8. Boyer, C. B., & Merzbach, U. C. (2011). *A History of Mathematics* (2nd ed.). John Wiley & Sons.
9. Chomsky, N. (1957). *Syntactic Structures*. Mouton.
10. Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press.
11. Grätzer, G. (2008). *Universal Algebra*. Springer.
12. Grohe, M. (2021). *The Logic of Artificial Intelligence*. Cambridge University Press.
13. Habash, N. (2010). *Introduction to Arabic Natural Language Processing*. Morgan & Claypool.

14. Hale, M., & Reiss, C. (2008). *The Phonological Enterprise*. Oxford University Press.
15. Hayes, B. (2009). *Introductory Phonology*. Wiley-Blackwell.
16. Ibn al-Bannāʾ al-Marrākushī. (ca. 1275). *Talkhīṣ Aʿmāl al-ḥisāb*.
17. Jurafsky, D., & Martin, J. H. (2023). *Speech and Language Processing* (3rd ed.). Pearson.
18. Kennedy, H. S. (1956). *The Development of Arabic Mathematics: Between Arithmetic and Algebra*. London.
19. Khan, M. A. (2020). *Mathematics in Islamic Tradition and Modernity*. Cambridge University Press.
20. Kornai, A. (2007). *Mathematical Linguistics*. Springer.
21. Marcus, S. (1967). *Algebraic Linguistics: Analytical Models*. Academic Press.
22. Marcus, G., & Davis, E. (2019). *Rebooting AI: Building Artificial Intelligence We Can Trust*. Pantheon Books.
23. Partee, B., ter Meulen, A., & Wall, R. (1990). *Mathematical Methods in Linguistics*. Kluwer.
24. Pollard, C., & Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. CSLI Publications.
25. Rashed, R. (1996). *A History of Arabic Mathematics: Between Arithmetic and Algebra* (A. F. Wafi, Trans.). Springer.
26. Rota, G.-C. (1997). *Indiscrete Thoughts*. Birkhäuser.
27. Ryding, K. C. (2005). *A Reference Grammar of Modern Standard Arabic*. Cambridge University Press.
28. Stewart, I. (2015). *Galois Theory* (4th ed.). Chapman and Hall/CRC.

29. Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5185–5198.
30. Bohas, G. (1990). *Les structures de la racine arabe*. Presses Universitaires de France.
31. Boudelaa, S., & Marslen-Wilson, W. D. (2001). Morphological units in the Arabic mental lexicon. *Cognition*, 81(1), 65–92. https://doi.org/10.1016/S0010-0277(01)00119-6
32. Mohri, M. (2009). Weighted automata algorithms. In M. Droste, W. Kuich, & H. Vogler (Eds.), *Handbook of Weighted Automata* (pp. 213–254). Springer.
33. Pullum, G. K., & Scholz, B. C. (2001). On the distinction between grammaticality and acceptability. *The Linguistic Review*, 18(2–4), 147–170.

رابعا: أوراق علمية وتقارير ما قبل النشر (arXiv & Technical Reports)

34. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *arXiv preprint*, arXiv:2005.14165.
35. Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint*, arXiv:1810.04805.
36. Marcus, G. (2020). The Next Decade in AI: Why Common Sense Matters. *arXiv preprint*, arXiv:2002.06177.

خامسا: المواقع الإلكترونية والمصادر المفتوحة (Web Resources)

37. Bender, E. M., & Koller, A. (2020). *Climbing towards NLU: On meaning, form, and understanding in the age of data*. ACL Anthology. https://www.aclweb.org/anthology/2020.acl-main.463
38. Jurafsky, D., & Martin, J. H. (2023). *Speech and Language Processing* (3rd ed. Draft). https://web.stanford.edu/~jurafsky/slp3/