الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي
جامعة أبي بكر بلقايد - تلمسان
كلية الآداب واللغات
قسم اللغة والأدب العربي

رسالة مقدمة لنيل شهادة الدكتوراه

في اللسانيات التطبيقية بعنوان:

توليد الجمل في اللسان العربي
(دراسة لسانية حاسوبية)

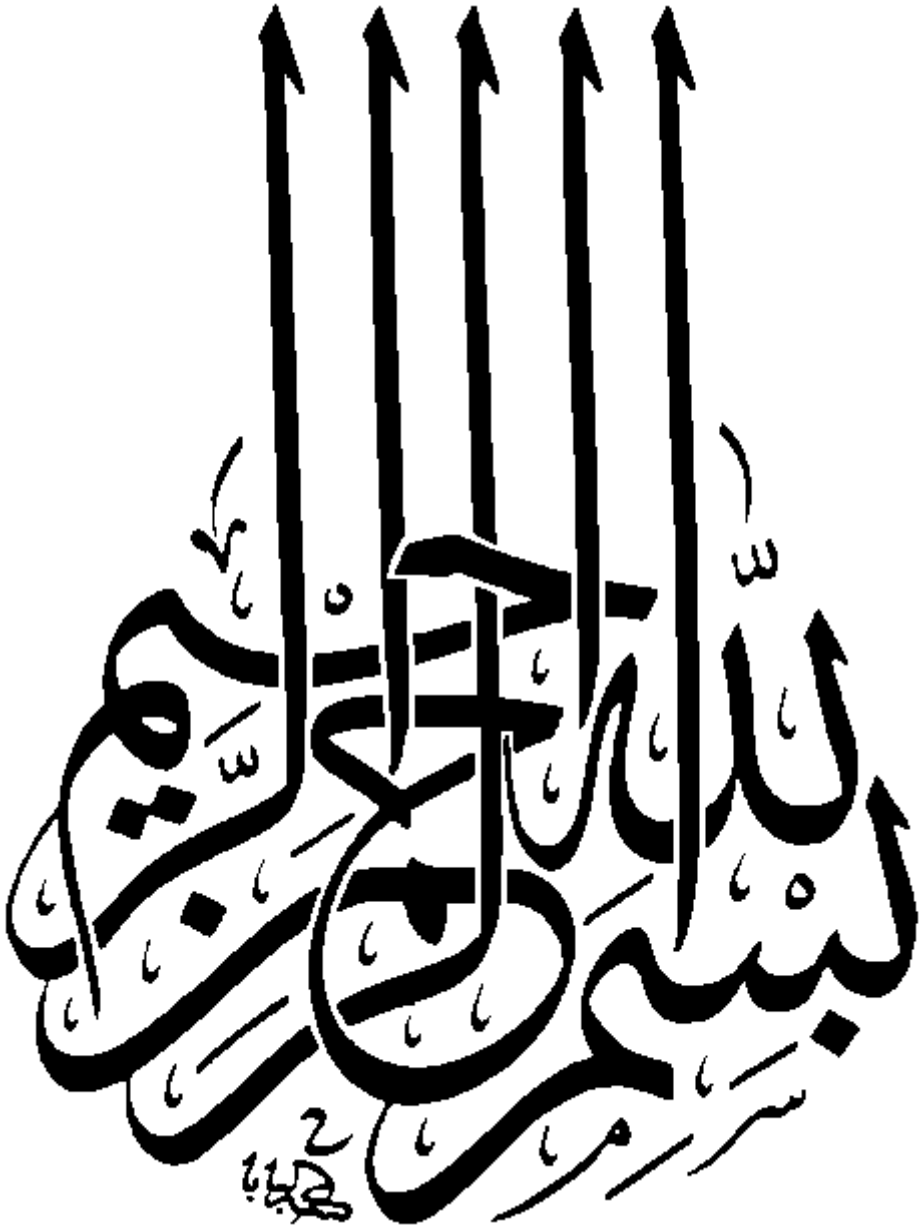
إعداد الطالب:
- بابا أحمد رضا

إشراف الأستاذ الدكتور:
- غيثري سيدي محمد

لجنة المناقشة:

- أ/د. ديدوح عمر	أستاذ التعليم العالي	جامعة تلمسان	رئيسا
- أ/د. غيثري سيدي محمد	أستاذ التعليم العالي	جامعة تلمسان	مشرفا ومقررا
- أ/د. سالمي عبد المجيد	أستاذ التعليم العالي	جامعة الجزائر	عضوا
- أ/د. التواتي بن التواتي	أستاذ محاضر (أ)	جامعة الأغواط	عضوا
- أ/د. بسناسي سعاد	أستاذ التعليم العالي	جامعة وهران	عضوا
- أ/د. عبد الرحيم محمد الأمين	أستاذ التعليم العالي	جامعة تلمسان	عضوا

السنة الجامعية 2014/2013


بسم الله الرحمن الرحيم

شـــكـــر وتـــقـــديـــر

﴿ الحَمْدُ للهِ الذِي هَدَانَا لِهَذَا وَمَا كُنَّا لِنَهْتَدِيَ لَوْلَا أَنْ هَدَانَا اللهُ ﴾

أتقدم بخالص شكري وتقديري وامتناني إلى أستاذي المشرف الدكتور سيدي محمد غيثري لإشرافه على هذه الرسالة.

كما أبسط شكري الجزيل بين يدي أعضاء اللجنة الموقرة التي تعمل على تقويم هذه الرسالة.

وأشكر أيضا كل من ساندني في إنجاز هذا البحث، ومدني بيد المساعدة من قريب أو بعيد.

إهـــــــــــــداء

إلى كوثر ومنال

مـــقـــدمـــة

﴿ الرَّحْمَنُ عَلَّمَ الْقُرْآنَ خَلَقَ الْإِنْسَانَ عَلَّمَهُ الْبَيَانَ ﴾

(الرحمن، 1-4)

لقد حققت البحوث الحاسوبية الحديثة العديد من الإنجازات في معالجتها للمعلومة بصفة عامة وللمعلومة اللغوية بصفة خاصة، وانتشرت البرامج والأنظمة التي تساعد مستخدم الحاسوب على تحقيق كثير من أهدافه والاستفادة من مختلف الخدمات المتاحة، والتي توفر له إمكانية القيام بالكثير من المهام اللغوية كمعالجة النصوص، وتصحيح الأخطاء النحوية والإملائية، كما تتيح له فرصة ترجمة نصه من لغة إلى أخرى، بالإضافة إلى الكثير من الخدمات عبر شبكة الإنترنت والبريد الإلكتروني وغيرها من البرامج التي تتضمن من قريب أو بعيد معالجة اللغة معالجة آلية.

إن تلك البرامج غالبا ما تزود بأنظمة لإنتاج اللغة أو تحليلها حيث تقوم بتوليد مختلف المعطيات اللغوية كالنصوص والفقرات والجمل والوحدات اللغوية انطلاقا من قواعد بيانات وتمثيلات لكل المعلومات الخاصة بها، كما تقوم بتحليل تلك المعطيات اللغوية المختلفة والتعرف على المعلومات المشفرة فيها. وقد حظيت الكثير من اللغات كالإنجليزية والفرنسية والعربية بدراسات معمقة لتصميم برامج حاسوبية متطورة من أجل تسهيل استفادة المستخدم منها في تأدية مختلف المهام والأنشطة اللغوية والتفاعلية.

الإشكالية

يمثل اللسان العربي[1] من خلال تنوعاته اللغوية واللهجية المتعددة الاستعمالات والوظائف ذا أهمية بالغة، ليس فقط على الأصعدة الثقافية والدينية والأدبية، ولكن حتى العلمية والاقتصادية والاجتماعية في عصر ثورة المعلومات، لذلك تعد حوسبته ومعالجته آليا رهانا وتحديا بالنسبة للباحثين من أجل تطوير برامج حاسوبية قادرة على معالجة اللسان العربي من خلال لغاته[2] المتعددة المكتوبة والمنطوقة، الأدبية والتواصلية، الفصيحة والعامية...

ترتكز البرامج الحاسوبية التي تعمل على توليد المعطيات اللسانية وتحليلها بطريقة آلية على دراسات تتناول نشاطين لسانيين نفسيين[3] هما التوليد والتحليل على التوالي؛ لأنهما لا يعزلان النشاط اللساني عن مرتكزاته المعرفية والنفسية، وهما مجالان اجتهد الباحثون في نمذجتهما وحوسبتهما من أجل تطوير برامج الترجمة الآلية خاصة ضمن مجال الذكاء الاصطناعي.

---

[1] إن اللسان العربي هو مفهوم قرآني؛ «إِنَّا أَنْزَلْنَاهُ قُرْآنًا عَرَبِيًّا لَعَلَّكُمْ تَعْقِلُونَ» (يوسف، 2)، «وَإِنَّهُ لَتَنْزِيلُ رَبِّ الْعَالَمِينَ نَزَلَ بِهِ الرُّوحُ الأَمِينُ عَلَى قَلْبِكَ لِتَكُونَ مِنَ الْمُنْذِرِينَ بِلِسَانٍ عَرَبِيٍّ مُبِينٍ» (الشعراء، 192–195)، «لِسَانُ الَّذِي يُلْحِدُونَ إِلَيْهِ أَعْجَمِيٌّ وَهَذَا لِسَانٌ عَرَبِيٌّ مُبِينٌ» (النحل، 103)، يتضمن هذا المفهوم أداءات لغوية متعددة بعضها وارد في القراءات القرآنية، وهي تشمل لغات العرب الفصيحة كما تشمل لغات العرب غير الفصيحة أو "غير المستحسنة في قراءة القرآن وإنشاد الشعر" حملا على تعبير سيبويه (ينظر: سيبويه: الكتاب، تحقيق: عبد السلام هارون، دار الجيل، بيروت، ط1، د.ت، ج. 4، ص. 432).

[2] يستخدم الباحثون المعاصرون غالبا اللغة مرادفة للسان، والمقصود باللغة هنا هي صور اللسان المتعددة الاستعمال والتوظيف.

[3] هما نشاطان لسانيان نفسيان نظرا لكونهما عمليتين تنطلقان من المعنى/النص وصولا إلى النص/المعنى، وبالتالي .بهما يربطان بين ظواهر ذات طبيعة ذ .يـ)الـ....ي .....كار...) و لأ هـ ذا لأ لأبيلأة لالألألألألأ وا) لألجمل...(.

وعليه، اخترت في رسالتي هذه أن أركز على نشاط التوليد في صلته بالمعالجة الآلية للغة، وأن أركز على جوانبه اللسانية الحاسوبية، وأن أخصص منها جانبا تطبيقيا لتوليد الجمل في اللسان العربي متجسدا في اللغة العربية المكتوبة، وبالتالي وسمت رسالتي بعنوان:

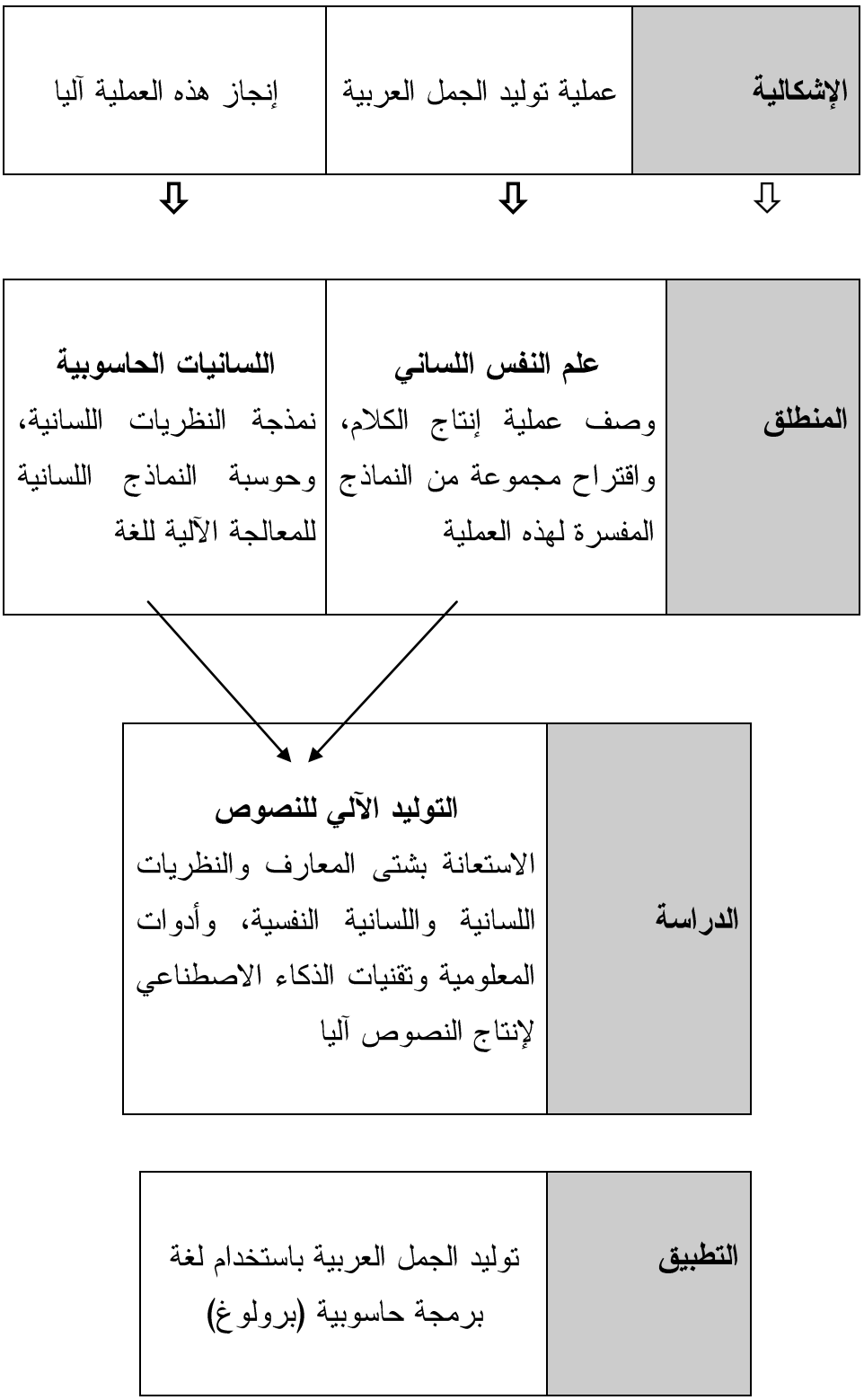
توليد الجمل في اللسان العربي

(دراسة لسانية حاسوبية)

إن هذا الموضوع يقتضي مني إثارة مجموعة من الأسئلة ضمن إشكالية عامة للبحث أحاول الاقتراب من الإجابة عنها عبر مكونات هذه الرسالة، وهي تتمثل فيما يلي:

- كيف يتمكّن المتكلم العربي من إنتاج الجمل وتوليدها ؟ وهل من الممكن إنجاز هذه العملية بطريقة آلية ؟

إن الاقتراب من هذه الإشكالية يقتضي الاستعانة بمجموعة من التخصصات وتداخلها هي النحو العربي الذي يدرس الجملة وأنواعها ومكوناتها، واللسانيات التي تدرس اللغة دراسة علمية، بالإضافة إلى اللسانيات النفسية التي تقترح نماذج لتوليد الجمل، وأخيرا اللسانيات الحاسوبية التي تعالج تلك النماذج معالجة آلية. وأفترض أن معالجة هذه الإشكالية تتم في خطوات متتالية ممثلة في المخطط التالي:



أسباب اختيار الموضوع

هناك عدة أسباب ذاتية وموضوعية دفعتني إلى اختيار هذا الموضوع. فلطالما جلبت انتباهي تلك الملكة التي يتمتع بها المتكلم في التعبير عن المعاني، وشدتني طريقة صياغته لتلك المعاني في جمل أو أقوال أو نصوص، كما أعجبتني قدرة الباحثين على وصف هذه العملية،

واختلافهم في تفسيرها وبيان مكوناتها، فشجعني ذلك على الاهتمام بدراسة توليد الجمل العربية بالتحديد لولعي الشديد باللغة العربية لغة الحضارة والدين والوطن.

ومن الأسباب الموضوعية التي جعلتني أفضل هذا الموضوع دون سواه، أن الدراسة الحاسوبية لتوليد الجمل ضرورية من أجل تصميم برامج في غاية الأهمية كبرامج الترجمة الآلية، والتي يحتاج تطويرها إلى وصف للنصوص والجمل العربية، وإلى توظيف أنظمة تنظيرية متماسكة من أجل صياغة نماذج التوليد.

بالإضافة إلى ذلك، إن الدراسة الحاسوبية لتوليد الجمل تفضي إلى تطبيق مجموعة من النظريات اللسانية واللسانية النفسية التي حاولت وصف اللغة، واقترحت لذلك فرضيات مختلفة حول اشتغالها من أجل تشجيعها على التفاعل مع مختلف النظريات الإعلامية والحاسوبية.

الأهداف

أتوخى من هذه الدراسة، وبالاستعانة بالمعلومات اللازمة، توظيف مجموعة من الأسس اللسانية الحاسوبية من أجل إقامة برنامج حاسوبي يمكنه توليد الجمل، وذلك من خلال تطبيق يتناول تكوين جمل عربية باستخدام لغة برولوغ (إحدى لغات البرمجة الحاسوبية)، وأركز في دراستي تلك على الجمل العربية المحددة في أشكالها المكتوبة، والبعيدة عن تحقيقاتها النطقية. وعليه، فإن لهذه الدراسة حدودا لا تتجاوزها، فليس بمقدور هذا البحث أن يقدم تصورا متكاملا لبرنامج حاسوبي متطور يتم عبره توليد كل الجمل في اللسان العربي انطلاقا من تمثيلاتها الدلالية، ولكن تسعى هذه الدراسة إلى تكوين جمل عربية بسيطة انطلاقا من قواعد بيانات تتضمن الكلمات وكيفية ائتلافها.

المنهج المتبع

عملت في هذا البحث على إضاءة المفاهيم الأساسية (كمفهومي الجملة والتوليد) التي يثيرها الموضوع، ثم على توظيفها في إطار المنهج الذي تتيحه الدراسة اللسانية الحاسوبية، والذي يقوم على وصف الظاهرة اللغوية وصورنتها من أجل تشغيلها آليا. وبالتالي، فقد انتهجت منهجا وصفيا وصوريا لعملية توليد الجملة العربية في الجانب النظري من الرسالة، ثم منهجا حاسوبيا في الجانب التطبيقي منها عند تنفيذ برامج التوليد اعتمادا على لغة برولوغ.

خطة البحث

يتضمن هذا البحث مقدمة منهجية وخمسة فصول تتعرض إلى المعالم النظرية التي يهتدي بها هذا البحث، وكذلك الجوانب العملية التي تبرز تطبيقاته. في الفصل الأول، تطرقت إلى بعض المفاهيم الضرورية للموضوع كالجملة واللسان مع التركيز على الجملة، فتناولت أبرز التعاريف التي مستها قديما وحديثا. وقد خصصت الفصل الثاني لأسس الدراسة اللسانية الحاسوبية نظرا لتعويلي عليها في مقاربة هذا الموضوع. وفي الفصل الثالث، تعرضت إلى توليد الجمل في مختلف النظريات اللسانية واللسانية النفسية، حتى يكون معتمدا للتوليد الآلي للجمل الذي سأتناوله في الفصل الرابع. وأخيرا وجب إشفاع هذه الدراسة بفصل تطبيقي يشتمل على برامج كتبت بلغة برولوغ لتكوين الجمل العربية تكوينا آليا. وقد تضمن الرسالة أيضا خاتمة تحتوي على أبرز نتائج الدراسة، وقائمة بالمصادر والمراجع المعتمدة، وقد أثبت في نهاية الرسالة مسردا للمصطلحات الواردة فيها مع ما يقابلها في اللغة الفرنسية والإنجليزية، وأخيرا فهارس للأشكال والجداول والموضوعات.

الصعوبات

أبرز الصعوبات التي لاقيتها في هذا البحث هو تعدد الأطر النظرية التي تناولت الجملة وتنوعها بل تعارضها أحيانا كما هي الحال في التعارض بين اللسانيات الصورية وبين اللسانيات الوظيفية أو المعرفية. بالإضافة إلى تعدد الاختصاصات التي وجب الاعتماد عليها كاللسانيات والمعلوميات وتداخلها وصعوبة الجمع بينها، بالإضافة إلى عدم تمكني من الحصول على بعض المراجع المهمة خاصة تلك التي تناولت معالجة اللغة العربية باستخدام لغة برولوغ.

أبرز المصادر والمراجع

استعنت في دراستي بمجموعة من المصادر والمراجع في نسخها الورقية المعهودة أو في نسخها الإلكترونية، وهي مصادر ومراجع باللغة العربية، أو باللغتين الفرنسية والإنجليزية بعد ترجمة نصوصهما إلى العربية، من أهمها:

- سيبويه (أبو بشر)، الكتاب، تحقيق: عبد السلام هارون.
- الحاج صالح (عبد الرحمن)، بحوث ودراسات في اللسانيات العربية.

- Bouillon (Pierrette) et *al*., Traitement automatique des langues naturelles.
- Danlos (Laurence), Génération automatique de textes en langues naturelles.
- Fayol (Michel), Production du langage.
- Fuchs (Catherine). et *al*., Linguistique et traitement automatique des langues.

وقد قمت، في أغلب الأحيان، بترجمة ذاتية للنصوص الفرنسية والإنجليزية التي رجعت إليها مرتكزا على بعض المعاجم العامة والمتخصصة مثل:

- إدريس (سهيل)، المنهل: قاموس فرنسي–عربي.

- معجم الحاسبات، منشورات مجمع اللغة العربية.

- Le Robert, Dictionnaire Français-Anglais, Anglais- Français.
- Morvan (P.), Ditt ionni ddddd’’’’’ ’’ m mmmm

الفصل الأول:

مفـــــاهيم أوليـــــة

الفصل الأول: مفاهيم أساسية

يتناول هذا البحث موضوعا محددا هو توليد الجمل وبالضبط في اللسان العربي، لذلك فهو يتخذ مسارا تطبيقيا؛ إذ لا يتناول عملية توليد الجملة تناولا لسانيا عاما بل يستدعي تلك القوانين والمعلومات المقررة في اللسانيات العامة حول توليد الجملة ليوظفها في دراسة عملية توليد الجملة العربية، وبعد ذلك قد يمد هذا البحث اللسانيات العامة بما يمكن أن يتوصل إليه من نتائج خاصة بتوليد الجملة العربية.

قبل ذلك، ينبغي أن أحدّد تلك المكونات المشكلة للموضوع أو بعبارة أخرى: ماذا يُقصد بـــ: *التوليد والجملة واللسان العربي* ؟ أحاول في هذا الفصل أن أتعرض لمفهومي *الجملة واللسان العربي*، وسأخصص فصلا مستقلا لمفهوم *التوليد*.

1. الجملة عند النحاة

إن التفكير في الجملة وما تحتوي عليه من عناصر وما تقابلها من دلالات تركيبية من أول ما شد انتباه الدراسين العرب، غير أني أبدأ بمفهوم الجملة عند سيبويه لأنه يمثل –مع شيخه الخليل– مرحلة صياغة المفاهيم اللغوية ونضج المنهج النحوي، وكتابه هو أول كتاب نحوي يصل إلينا.

1.1. الجملة عند سيبويه

إن المتتبع لما ورد في كتب النحو العربي يجد اختلافا بين النحاة في تحديد الجملة، بل يجد تضاربا حول اتخاذها مصطلحا للدلالة على مفهوم علمي معين. في أقدم وأكمل كتاب

نحوي وصل إلينا، نعني كتاب سيبويه، نلفيه يستخدم في كثير من المواضع مصطلحا قريبا من مصطلح الجملة وهو الكلام المستغني.

من خلال تحليلنا للنص الآتي –في باب ما ينتصب فيه الخبر– نحاول أن نقف على مدلول الكلام المستغني، يقول سيبويه: « ألا ترى أنك لو قلت: فيها عبد الله، حسن السكوت وكان كلاما مستقيما[1]، كما حسن واستُغني في قولك: هذا عبد الله، وتقول: عبد الله فيها، فيصير كقولك: عبد الله أخوك، إلا أن 'عبد الله' يرتفع مقدما كان أو مؤخرا بالابتداء. ويدلك على ذلك أنك تقول: إن فيها زيدا، فيصير بمنزلة قولك: إن زيدا فيها، لأن 'فيها' لما صارت مستقرا لزيد يستغني به السكوت وقع موقعَ الأسماء، كما أن قولك: عبد الله لقيته، يصير 'لقيته' فيه بمنزلة الاسم، كأنك قلت: عبد الله منطلق، فصار قولك 'فيها' كقولك: استقر عبد الله، ثم أردت أن تخبر على أية حال استقر فقلت: قائما، فـ'قائم' حال مستقر فيها، وإن شئت ألغيت 'فيها' فقلت: فيها عبد الله قائم...»[2]

يقصد سيبويه بالكلام المستغني تلك العبارة القولية[3] التي يحسن أن يسكت المتكلم في آخرها وينقطع عن لفظ آخر قد يرد بعدها، ولم ينتظر المخاطَب معنى إضافيا لتكتمل لديه المعلومة، ويتضح ذلك من خلال مثال: فيها عبد الله قائم أو قائما، فالرفع يدل على أنه في هذه

---

[1] الكلام المستقيم عنده هو قسيم المحال، ويقصد به السليم من جهة تركيبه اللغوي وهو بدوره ينقسم إلى حسن وقبيح وكذب بحسب درجة مقبوليته لدى المتكلمين. ينظر: سيبويه (أبو بشر)، الكتاب، ت. عبد السلام هارون، دار الجيل، بيروت، ط. 1، د.ت، ج. 1، ص. 25–26.

[2] نفسه، ج. 2، ص. 88–90.

[3] قد فصل ابن جني الحديث في الفرق بين الكلام والقول ملخصه أن كلا منهما لفظ يتضمن معنى إلا أن الكلام خاص بما توافرت فيه صفتا الاستقلال والإفادة. ينظر: ابن جني (أبو الفتح)، الخصائص، ت. محمد علي النجار، المكتبة العلمية، د.ط، د.ت، ج. 1، ص. 77.

العبارة لا يمكن أن نستغني عن 'قائم' دون أن يؤثر ذلك في مضمون الكلام ومراده لأن المقصود هو الإخبار بقيام عبد الله لذلك أمكن أن يشكل الجميع كلاما واحدا، أما النصب فيدل على إمكانية الاستغناء عن 'قائم' لأن المقصود في هذه الحال هو الإخبار بمكان 'عبد الله'، وفي كلتا الحالتين نلاحظ مقصد الإفادة وكذا انفصال العبارة عن غيرها بسكوت يحدث لدى المخاطب اكتفاء واستيفاء للمعلومة. وبالتالي فإن الكلام المستغني عند سيبويه « يشكل وحدة تبليغية تتم بها الفائدة للمخاطب»[1].

الملاحظ من هذا التحديد وجوب توافر ثلاثة عناصر حتى يتم مفهوم الكلام المستغني هي:

- قصد المتكلم الإخبار بشيء معين، مما يجعل للعبارة وظيفة تعبيرية؛
- حصول العلم بشيء معين أو تحقيق الفائدة عند المخاطَب، وليس العلم هنا أن يكون للعبارة معنى، ولكن أن يقع تأثير المتكلم في السامع، مما يجعل للعبارة وظيفة اتصالية؛
- ظهور العبارة القولية في سلسلة صوتية واحدة بين سكوتين: سكوت يسبق الكلام وسكوت يتبعه. ويتعلق حسن السكوت بمدى حصول المعلومة لدى المخاطب، مما يدل على أن هناك علاقة طردية بينهما فكلما حصلت المعلومة حسن السكوت.

---

[1] الحاج صالح (عبد الرحمن)، بحوث ودراسات في اللسانيات العربية، موفم للنشر، الجزائر، د.ط، 2007، ج. 1، ص. 291.

	المتكلم	المخاطب
الكلام	استغناء	استحسان
القصد	إخبار	إعلام

الجدول1: الكلام المستغني عند سيبويه

بالمقابل، هل يمكن أن نستخلص من هذا ومن أمثلة سيبويه بأن الإعلام أو الإفادة متصلان بالكلام الخبري فقط، وبالتالي فإن العبارة الإنشائية لا يطلق عليها مصطلح "الكلام" خاصة وأنه يقول في موطن آخر –حيث يتناول إمكانية نصب 'كنتُ' لجملة اسمية–: « ألا ترى أنه لم ينفِذ الفعل في 'كنتُ' إلى المفعول الذي به يستغني الكلام كاستغناء 'كنت' بمفعوله، فإنما هذه في موضع الإخبار وبها يستغني الكلام. »[1] فجعل الكلام المستغني هنا قرينا للإخبار ؟

الإجابة عن ذلك: إن الإعلام أو الإفادة ليسا متصلين بالكلام الخبري فقط لأن شرط الإفادة – كما سيبينه ابن يعيش فيما بعد – أن توجد هناك علاقة الإسناد وهي نسبة تشمل الخبر وغيره من الأساليب البلاغية كالأمر والنهي والاستفهام، فالهدف إذن هو تحقيق معلومة لدى المخاطَب سواء أكانت إخبارا بموضوع ما أو داعية إلى فعل ما[2].

2.1. أولية المصطلح

أما عن مصطلح الجملة فيرجح عبد الرحمن الحاج صالح بأن الأخفش تلميذ سيبويه وراوية كتابه وأستاذ المازني قد يكون أول من ابتدع مصطلح الجملة مقابلا للكلام المستغني

---

[1] سيبويه (أبو بشر)، الكتاب، مرجع سابق، ج. 1، ص. 149.

[2] ابن يعيش (موفق الدين)، شرح المفصل، ت. أحمد السيد أحمد، المكتبة التوفيقية، القاهرة، د.ط، د.ت، مجلد 1، ج. 1، ص. 44.

لأنه أشار في كتابه "القوافي" إلى أبرز مكون لمفهوم الجملة ألا وهو الفائدة، ويظن أنه انتقل عبر المازني إلى المبرد[1].

بالمقابل، من المسلم به أن أول ظهور لمصطلح الجملة كان في كتاب "المقتضب" للمبرد حيث يقول: «وإنما كان الفاعل رفعا لأنه هو والفعل جملة يحسن عليها السكوت، وتجب بها الفائدة للمخاطب، فالفاعل والفعل بمنزلة الابتداء والخبر، إذا قلت: قام زيد بمنزلة قولك: القائم زيد. »[2]

إن المبرد –بهذا القول– يجعل الجملة مرادفة للكلام المستغني عند سيبويه، ويظهر ذلك بشكل أوضح حينما يقول في موطن آخر: « فالابتداء نحو قولك: زيد، فإذا ذكرته فإنما تذكره للسامع ليتوقع ما تخبره به عنه، فإذا قلت: منطلق أو ما أشبهه صح معنى الكلام، وكانت الفائدة للسامع في الخبر »[3]. يبدو من خلال هذا النص اهتمام هذا النحوي بالعناصر المقررة سلفا وهي المخبر والسامع وحصول الفائدة في عملية الإخبار، مما يجعلنا نستنتج: *أن الجملة هي أصغر وحدة لغوية تحققت فيها الفائدة أو المعلومة*.

3.1. الجملة عند ابن فارس وابن جني

من أبرز من عرف الكلام أيضا ابن فارس إذ يقول: « زعم قوم أن الكلام ما سُمع وفُهم، وذلك قولنا: قام زيد وذهب عمرو، وقال قوم: الكلام حروف مؤلَّفة دالة على معنى. » ثم عقَّب على هذين التعريفين قائلا: « والقولان عندنا متقاربان؛ لأن المسموع المفهوم لا يكاد

---

[1] الحاج صالح (عبد الرحمن)، بحوث ودراسات في اللسانيات العربية، مرجع سابق، ج. 1، ص. 291.

[2] المبرد (أبو العباس)، المقتضب، ت. محمد عبد الخالق، عالم الكتب، بيروت، د.ط، ج. 1، ص. 8.

[3] المبرد (أبو العباس)، المقتضب، مرجع سابق، ج. 4، ص. 126.

يكون إلا بحروف مؤلفة تدل على معنى. »[1] صحيح أن القولين متقاربان غير أنهما ينطلقان من موقفين متباينين: الأول موقف المتلقي والآخر موقف المرسل، أو بعبارة أخرى: الحال الأولى هي حال المخاطب الذي يتلقى عبر حاسة السمع ثم يترجم ما يتلقاه ويحلله ويفهمه، أما الحال الثانية فهي حال المتكلم الذي يصوغ معانيه الذهنية في حروف ومقاطع صوتية.

الحال	الكلام
المتلقي	هو الملتقط سماعا، ويتفرع عنه المكتوب الذي يلتقط بالبصر ولا يتفرع عنه غيره مما يدرك بسائر الحواس (الميزة الصوتية للكلام)
المرسل	هو المؤلف من مقاطع ذات معنى (ارتباط تلك المقاطع وائتلافها بما يضمن تحقيق الوحدة المعنوية)

الجدول2: تصور الكلام عند ابن فارس

يبدو من تمثيل ابن فارس أن الكلام عنده مرادف للجملة، وهذا ما نلمسه عند معاصره ابن جني الذي يعرف الكلام بأنه « كل لفظ مستقل بنفسه مفيد لمعناه وهو الذي يسميه النحويون الجمل نحو: زيد أخوك وقام محمد وضُرب سعيد وفي الدار أبوك وصه ومه ورويد وحاءِ وعاءِ في الأصوات وحسِّ ولبِّ وأف وأوه، فكل لفظ استقل بنفسه وجنيت منه ثمرة معناه فهو كلام. »[2] تلفت انتباهنا من هذا التعريف مجموعة من الأشياء:

---

[1] ابن فارس (أبو الحسين)، الصاحبي في فقه اللغة العربية ومسائلها وسنن العرب في كلامها، ت. عمر فاروق الطباع، مكتبة المعارف، بيروت، ط.1، 3993، ص81.

[2] ابن جني (أبو الفتح)، الخصائص، مرجع سابق، ج. 1، ص. 77.

- استعمال ابن جني للكلام مفردا في مقابل الجمل التي وردت في صيغة الجمع، ويمكن أن نفسر ذلك بأن الكلام مصدر (أو بالإحرى هو اسم مصدر للفعل "تكلم")، والمصدر يدل على الحدث –كما يقول النحاة–، والحدث لا يتعدد، بخلاف الجملة التي هي اسم جنس فيمكن أن تثنى وتجمع.
- الكلام والجمل كل منهما يتحقق فيه استقلال اللفظ مع إفادة المعنى.
- الجمل مصطلح نحوي والكلام مصطلح شائع في عدة علوم كعلم الكلام وأصول الفقه.

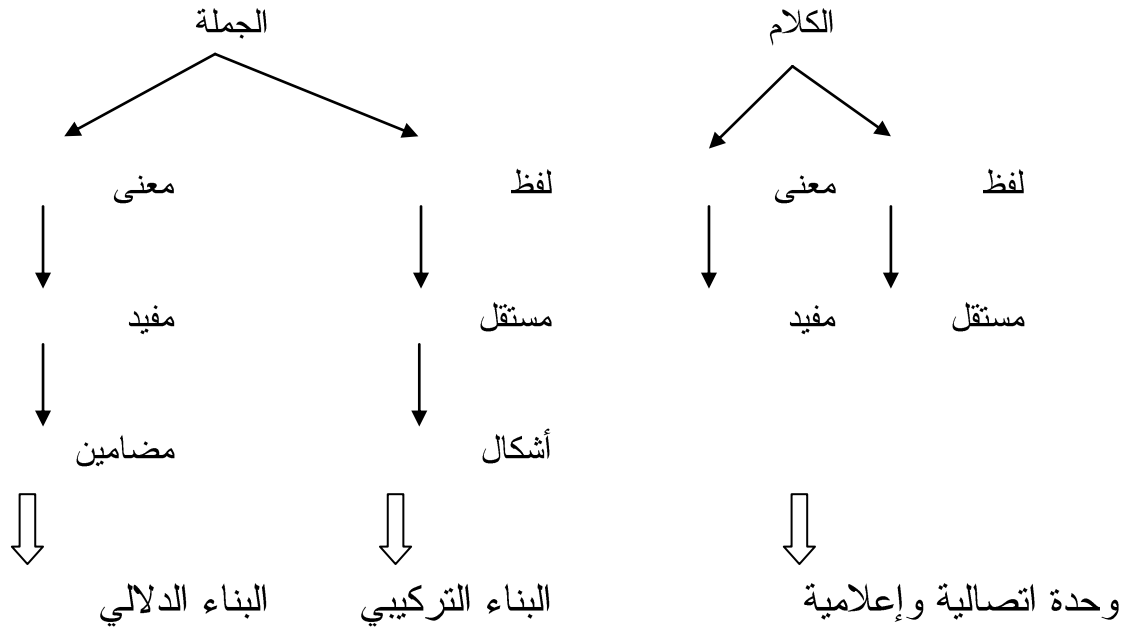
نستنتج من ذلك: أن السبب في إفراد الكلام هو نظر النحاة إليه على أنه وحدة لفظية ومعنوية، مهما زادت عناصر مجموعة الألفاظ والمعاني المشكلة للكلام أو قلّت تلك العناصر، ومهما كثرت جمله أو قلَّت، وبالتالي فإن الكلام هو واحد في ذاته وإن تغيرت أشكاله ومضامينه طالما بقيت فيه صفتا الاستقلال والإفادة، وهما صفتان تتعلق أولاهما بالبناء التركيبي والدلالي للأقوال والتعابير، وتتعلق الثانية بالوظيفة الإعلامية والاتصالية التي يؤديها الكلام.

أما الجمل فهي موضوع الدراسة النحوية حيث الاهتمام بسلامة البناء التركيبي والدلالي للملفوظات ليس إلا، وهذا الذي يؤكده ابن جني في موضع آخر: « الكلام إنما هو في لغة العرب عبارة عن الألفاظ القائمة برؤوسها المستغنية عن غيرها، وهي التي يسميها أهل هذه الصناعة الجمل على اختلاف تراكيبها. »[1] فبين أن الجمل تتعدد أنواعها وتختلف طرق تأليف عناصرها، وذلك واضح من خلال الأمثلة التي أوردها:

---

[1] ابن جني (أبو الفتح)، الخصائص، مرجع سابق، ج. 1، ص. 32.

- زيد أخوك – في الدار أبوك: جملتان اسميتان الأولى خبرها اسم والثانية خبرها ظرف مستقر قد تقدم على المبتدأ.
- قام محمد – ضُرب سعيد: جملتان فعليتان الأولى فعلها مبني لما يسمى فاعله والثانية فعلها مبني لما لم يسمَّ فاعله.
- صه ومه وروید وحسِّ ولبِّ وأف وأوه: جمل مكونة من أسماء الأفعال وضمائر الفاعلين المقدرة، وقد اختلفت أشكال أسماء الأفعال بين جامد ومشتق، وبين ما يدل على أمر أو مضارع.
- حاءِ وعاءِ: جملتان مكونتان من اسم صوت وفاعله.



الشكل1: الكلام والجملة عند ابن جني

إن التعريف الذي قرره ابن جني قد أقره من جاء بعده، فنجد ابن يعيش مثلا يشير إلى أن « الكلام –عند النحويين– عبارة عن كل لفظ مستقل بنفسه مفيد لمعناه ويسمى الجملة»[1]. ثم يوضح بأن الكلام الذي تحصل منه الفائدة هو الذي ينعقد بتركيب معين، فيشمل ذلك التركيب اسمين (المبتدأ وخبره) أو فعلا واسما (الفعل وفاعله) قد قام بينهما تعلق إسنادي، وهي تلك العلاقة التي تربط على المستوى الدلالي بين خبر ومخبر عنه أو بين حديث ومحدث عنه أو بصفة عامة: بين مسند ومسند إليه حتى يشمل الكلام الإنشائي.

ويبدو التفريق بين الكلام والجملة أكثر وضوحا عند ابن هشام الأنصاري حيث يعرف الكلام بأنه القول المقصود المفيد الذي يحسن السكوت عليه، أما الجملة فهي عبارة عن الفعل وفاعله، والمبتدأ والخبر، وما كان بمنزلة أحدهما، فلا يشترط في الجملة أن تكون مفيدة، والدليل على ذلك قول النحاة: جملة الشرط، جواب الشرط، جملة الصلة، وهذه كلها ليست كلاما لأنها ليست مفيدة[2].

وعليه، فإذا كانت الكلمات تكشف عن تلك العوالم التي هي المعاني فإن التركيب بين الكلمات يحدد تلك العوالم ويوجهها، والعلاقة الإسنادية التي تربط بين المسند والمسند إليه تحدد الخطاب وتعرف الملفوظ لدى المخاطب، أما في غياب هذه العلاقة فيبقى للكلمات بعدها الرمزي والإيحائي ولكنها تفقد طابعها التواصلي الوظيفي.

---

[1] ابن يعيش (موفق الدين)، شرح المفصل، مرجع سابق، مجلد 1، ج. 1، ص. 44–45.

[2] ينظر: الأنصاري (ابن هشام)، مغني اللبيب عن كتب الأعاريب، ت. محمد محيي الدين عبد الحميد، المكتبة العصرية، بيروت، 1995، ج. 2، ص. 411.

2. الجملة في الدراسات الحديثة

اعتنت الدراسات اللسانية الحديثة اعتناء كبيرا بتحليل اللغة من أجل استخراج مكوناتها في مختلف المستويات اللسانية، كما خصت الجملة بحيز كبير من تلك العناية، وأقامت عليها فرضيات ونظريات متعددة سواء في ذلك عند الباحثين العرب الذين يملكون تراثا لسانيا عريقا أو لدى الباحثين الغربيين الذين اجتهدوا في تجديد البحث اللساني.

1.2. الجملة عند الدارسين العرب

يتعدد مفهوم الجملة عند علماء اللغة العربية المحدثين بسبب انتمائهم إلى مدارس ومذاهب لغوية شتى، بعضها متأثر بالدراسات العربية في أطوارها المتتالية، وبعضها متأثر بنظريات لغوية غربية، وبعضها مازج بين القبيلين، وقد يترتب عن ذلك اختلاف القواعد والأحكام اللغوية الخاصة بالجملة، فهناك من اللغويين العرب من يرى أن الجملة قول مركب مفيد دال على معنى يحسن السكوت عليه، فلا يخرجون بذلك عما درج عليه النحاة عامة والمتأخرون على وجه الخصوص.

ويرى بعضهم بأن الجملة تركيب إسنادي يفيد فائدة يحسن السكوت عليها، والغاية منها الاتصال والتفاهم بين أعضاء الجماعة اللغوية، وهي مجموعة ذات عناصر لغوية إسنادية، وقد أنشئت قصد التفاهم في بيئة لغوية. فيبين بذلك أهم خصائص الجملة وهي أنها شكل لغوي، تربط بين عناصره علاقة الإسناد، ولكنه ربط هذا الشكل بمحتوى يحسن السكوت عليه وليس هذا بمطرد لأن هناك من الجمل ما يكون ناقصا من حيث الإفادة، ولا ينفي ذلك كونها جملا.

من جهته، يخلص **الدجني** إلى أن التعريفين اللغوي والاصطلاحي للجملة يشيران إلى الجمع بين الكلمات، وإلى أن مدلولها يختلف عن مدلول الكلام[1]. إن هذا الرأي يتلاقى مع فرضية الفصل بين التركيب والدلالة عند توليد الجمل وفهمها بحيث أن تركيب الكلام أو شكله وهو الذي يسمى الجملة يدل على معاني نحوية مستقلة عن المعاني الدلالية والمعجمية.

وانطلاقا من النظرية اللسانية التي تعتبر اللغة نظاما كليا مكونا من أنظمة صغرى مترابطة ضمن مستويات متراتبة، يتشكل كل مستوى منها من وحدات لسانية، يرى **المسدي** أن الجملة هي الوحدة النحوية الكبرى في الكلام، وهي تتميز بترابط أجزائها واستقلاليتها، فلا يمكن أن تتدرج في بناء نحوي أوسع.

ومن جهة أخرى، ينبه **إبراهيم أنيس** بداية إلى الخلط الذي يمكن أن يقع بين مدلولي الجملة والقضية المنطقية التي تحتوي على موضوع ومحمول، والتي تمكن من استخراج الحكم المستفاد من ارتباط أحدهما بالآخر. أما الجملة فيجب أن تلتمس –حسب رأيه– من استعمال الناس ومما تواضعوا على استقلاله بالمعنى في كل كلام.

وعليه، فهو يعرف الجملة بأنها « في أقصر صورها أقل قدر من الكلام يفيد السامع معنى مستقلا بنفسه، سواء تركب هذا القدر من كلمة واحدة أو أكثر»[2]. ومما يلاحظ أن هذا التعريف يجيز أن تتركب الجملة من كلمة واحدة، أي إن فكرة الإسناد ليست لازمة لتشكيل جملة صحيحة.

---

[1] الدجني (فتحي عبد الفتاح)، الجملة النحوية نشأة وتطورا وإعرابا، مكتبة الفلاح، الكويت، ط. 2، 7987، ص. 18، 30–32.

[2] أنيس (إبراهيم)، من أسرار اللغة، مكتبة الأنجلو المصرية، القاهرة، ط. 6، 7978، ص. 277.

ينبه **عبد الرحمن أيوب** بدوره أيضا إلى الفرق بين الجملة باعتبارها أمرا واقعيا، وبينها باعتبارها نموذجا تصاغ الجمل الواقعية على قياسه، وقد ضرب لذلك مثلا بقوله: "محمد قائم"، فهذا القول يصف نموذج الجملة الاسمية وهو تركيبها من مبتدإ وخبر، كما أنه يجسد نموذج الجملة الاسمية السابق في الألفاظ المستعملة، أو بعبارة أخرى يجدر التفريق بين الجملة باعتبارها مادة الكلام، وبين الجملة باعتبارها القالب الذي تصاغ فيه تلك المادة[1].

أما **ميشال زكرياء** فيرى في تعريف القدماء للجملة بأنها اللفظ المفيد فائدة يحسن السكوت عليها، هو تعريف دقيق يتبناه في دراسته، وهو يتلاقى مع تعريف اللسانيين المعاصرين للجملة بأنها وحدة كلامية مستقلة بالإمكان ملاحظتها عبر السكوت الذي يحددها[2].

ومن النظرية اللسانية العامة ينطلق **المخزومي** في دراسته للجملة فيعرفها بأنها: « الصورة اللفظية الصغرى للكلام المفيد في أية لغة من اللغات، وهي المركب الذي يبين المتكلم به أن صورة ذهنية كانت قد تألفت أجزاؤها في ذهنه ». يظهر من هذا التعريف بأن **المخزومي** لا يعتبر الجملة ظاهرة خاصة باللغة العربية، ولكنها ظاهرة لغوية عامة مشتركة بين جميع اللغات، كما يعتبرها وسيلة لترجمة الصورة المعنوية التي تشكلت في ذهن المتكلم[3].

ثم يضيف الباحث المذكور مبينا بعد ذلك عناصر الجملة قائلا: « والجملة التامة التي تعبر عن أبسط الصور الذهنية التامة التي يصح السكوت عليها تتألف من ثلاثة عناصر رئيسية

---

[1] نحلة (محمود أحمد)، مدخل إلى دراسة الجملة العربية، دار النهضة العربية، بيروت، 1988، ص. 77.

[2] ينظر: زكريا (ميشال)، الألسنية التوليدية والتحويلية وقواعد اللغة العربية، المؤسسة الجامعية، بيروت، ط. 2، 1986، ص. 24.

[3] المخزومي (مهدي)، في النحو العربي، نقد و توجيه، دار الرائد العربي، بيروت، ط. 2، 1986، ص. 11.

هي: المسند إليه أو المتحدث عنه أو المبني عليه والمسند الذي يبني على المسند إليه ويتحدث به عنه، والإسناد أو ارتباط المسند بالمسند إليه »[1].

يشير هذا التعريف إلى أن هناك ثلاث مكونات يتألف منها أي تركيب حتى يشكل جملة تامة، وهي المسند إليه والمسند وعلاقة الإسناد، وقد ذكر للمسند إليه مرادفين هما المتحدث عنه والمبني عليه. إن هذه المصطلحات وإن كان بينها ترادف على سبيل التوسع، لكن لكل منها توظيف خاص عند النحاة والبلاغيين؛ فاستخدام المسند إليه والمتحدث عنه غالبا ما يكون في سياق الحديث عن مضمون الجملة ومحتواها الدلالي والاتصالي، لذلك هو فمبحث من مباحث البلاغة والمعاني، على عكس المبني عليه الذي يتصل بشكل الحديث فهو مبحث من مباحث النحو.

وهناك فريق من الباحثين العرب لا يعتبر التركيب بين المسند والمسند إليه شرطا لتكوين الجملة؛ إذ يكفي أن تحتوي الجملة على كلمة واحدة ما دامت الفائدة متوفرة فيها[2]. لا يمكن التسليم بهذا الرأي لأن ذلك محمول عند النحاة على حذف عناصر أخرى والاستغناء عنها لوجود قرينة دالة عليها كالسياق مثلا، وهذه العناصر المحذوفة كأنها مذكرورة لكن استعيض عنها بدلالة السياق.

ضمن هذه الفرضية الصورية لبنائها، يرى **عبد الرحمن الحاج صالح** في الجملة رأيا أصيلا مفاده أن الجملة هي مثال الكلام وبناؤه كما أن للكلمة بناء ومثالا، أي إنه إذا أمكن تجريد

---

[1] المخزومي (مهدي)، في النحو العربي، نقد و توجيه، مرجع سابق، ص. 11.

[2] ينظر: عبد اللطيف (محمد حماسة)، العلامة الإعرابية في الجملة بين القديم والحديث، دار غريب، القاهرة، د.ط، 2001، ص. 57.

الكلمات إلى أوزان وصيغ وإبدال أحرف ثابتة (الفاء والعين واللام) بالحروف الأصلية لتلك الكلمات حتى تظهر صيغها، فكذلك يمكن أن يجرد الكلام إلى فعل وفاعل ومبتدأ وخبر.

بل أكثر من ذلك؛ لقد لاحظ هذا الباحث بأن بنية الفعل والفاعل والمفعول به من جهة وبنية المبتدأ والخبر من جهة ثانية متكافئتان، وبالتالي يمكن أن يوغل في مستوى أعلى من التجريد فيجعل الفعل في موضع العامل والفاعل في موضع المعمول الأول والمفعول به في موضع المعمول الثاني، والمبتدأ والخبر في موضعي المعمولين الأول والثاني على التوالي، بينما العامل فيهما هو الابتداء وهو بمنزلة الصفر أي الخلو من العامل اللفظي.

وأشار هذا الباحث أيضا إلى أن عناصر كثيرة قد تقوم مقام الفعل والابتداء والجامع بينها هو كونها تعمل في معمولين لفظا ومعنى، فيكون العامل محور الجملة، وأما باقي الفضلات من مفاعيل غير المفعول به وسائر المنصوبات (الحال والتمييز والمستثنى) فمخصصات داخل الجملة. وأما الزوائد التي تلحق بالاسم (كحروف الجر والتوابع) أو بالفعل (كحروف النصب والجزم) فتكوّن –مع الرأس اسما كان أو فعلا– ما يسميه **الحاج صالح** بلفظتي الاسم والفعل، وهاتان اللفظتان هما بمنزلة الاسم والفعل أين تكون عوامل أو معمولات.

كما تبين لهذا الباحث –من خلال تتبعه لآراء الخليل وسيبويه– عدة أحكام تتعلق ببنية العامل ومعموليه أوردها فيما يلي:

- إن العامل ومعموله الأول الذي يشغل به هما أقل ما تكون عليه الجملة الفعلية في اللغة العربية التي لم يقع فيها حذف، إذ لو وقع حذف أحدهما ينوى أنه مذكور في الجملة،

وهما في هذه الحالة الفعل والفاعل. وأما في الجملة الاسمية فالمعمول الثاني (الخبر) أيضا لازم الذكر أو مقدر إذا حذف.

- لا يقدم المعمول الأول على عامله البتة، وعلى ذلك فهما يكونان رياضيا زوجا مرتبا.
- قد يتبوأ موضع العامل حرف كـــ"إن" أو فعل ناسخ، كما أننا قد نجد في هذا الموضع عاملا مع معموله، أو عاملا مع معموليه (كما هو ملاحظ في الجدول3).
- إن موضع المعمول في الأصل معدّ للفظة الاسم (الاسم وما يدخل عليه من زوائد تخصه)، وقد تحل محلها جملة أو تركيب.

العامل	المعمول1	المعمول2
قرأ	الطالب	الكتاب
Ø	الكلمة الطيبة	صدقة
لعل	الفرج	قريب
حسبت	الخبر	صادقا
أعلمت صديقي	البريد	واصلا

الجدول3: تكافؤ البنى التركيبية العربية

ويضاف إلى ذلك أن الجملة التي تتكون من العامل ومعموليه هي البنية اللفظية للكلام، أما من الجانب الوظيفي الدلالي فيتكون الكلام من مسند ومسند إليه. وتجدر الإشارة هنا إلى أن

بعض الباحثين يسمي الجملة التي تحتوي على مسند ومسند إليه بالجملة الأساسية، وأما الجملة التي أضيفت إليها عناصر أخرى إلى جانب المسند والمسند إليه، فتدعى بالجملة المولَّدة[1].

وقد مثل الحاج صالح للبنية اللفظية أو الجملة بالنموذج الآتي:

[(ع ← م1) ± م2] ± خ

حيث:

ع: العامل؛

م1: المعمول الأول؛

م2: المعمول الثاني؛

خ: المخصص؛

(ع ← م1): زوج مرتب أو مركب

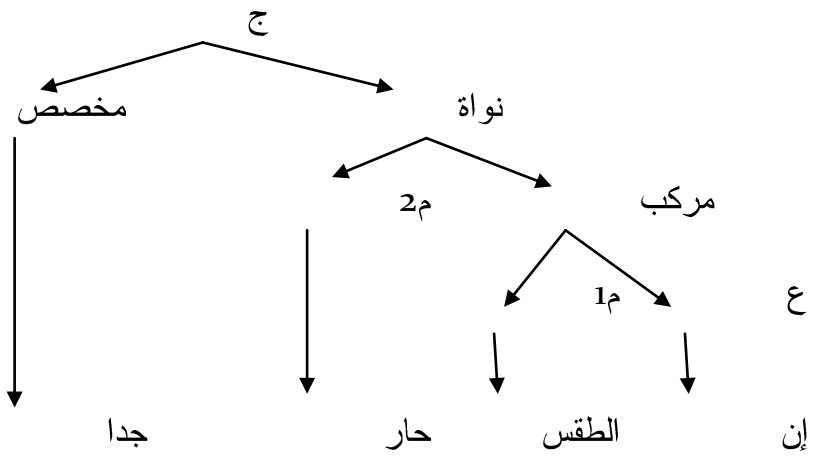
[(ع ← م1) ± م2]: النواة[2]

يمكن أن أمثل تبعا لذلك النموذج بالمشجر التركيبي لجملة مثل: 'إن الطقس حار جدا'، والغرض من ذلك الوقوف على المكونات المباشرة لهذه الجملة.

---

[1] ينظر: منصوري (ميلود)، دلالات التراكيب في نحو الجملة، دار أم الكتاب، مستغانم، ط. 1، 2033، ص. 20، 205.

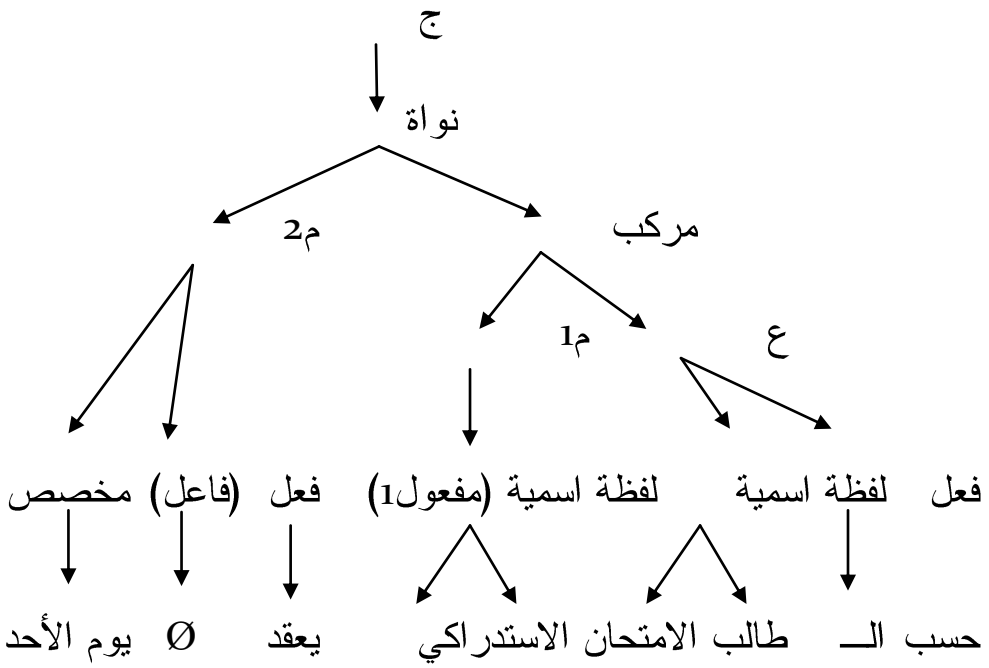
[2] ينظر: الحاج صالح (عبد الرحمن)، بحوث ودراسات في اللسانيات العربية، مرجع سابق، ج. 2، ص. 33-75.



وعليه أستنتج بأن البناء الأصلي للجمل في اللسان العربي يكون إما على شكل جملة فعلية أو جملة اسمية.

- الجملة الفعلية:
  - فعل + فاعل + مفعول به + مخصص
- الجملة الاسمية:
  - ع = Ø: مبتدأ + خبر + مخصص
  - ع = حرف ناسخ: حرف ناسخ + اسم + خبر + مخصص

هذه المواضع تحتوي على كلمات من نوع الأسماء المتكمنة، لكن باستخدام مبدأ دمج المكونات بعضها في بعض نحصل على مكونات شتى مثلا: 'حسب الطالب الامتحان الاستدراكي يعقد يوم الأحد'.



ملحوظة: في هذا المثال، المفعول الثاني لـــ"حسب" هو جملة فعلية لذلك فإن هذا المثال يتكون من جملة كبرى –على حد تعبير ابن هشام الأنصاري–[1]، كما أنه في هذا المستوى التركيبي ليس هناك مخصص للجملة لذلك لم يمثَّل في المشجر، لكن يوجد مخصص في مستوى الجملة التي حلت محل المفعول به الثاني.

بقي أن نشير إلى أن تغيير المواقع بالنسبة لمكونات الجملة –فيما تجوزه قواعد اللسان العربي طبعا– لا يؤثر في بنيتها مثلا: على حينِ غرةٍ اقتحمَ القلعةَ العدوُّ.

وعلى العموم، اختلف علماء اللغة في مسائل شتى حول الجملة النحوية وتحديد مفهومها وتسميتها وأقسامها وأنواعها، إلا أنهم اتفقوا على أن الجملة لا تتألف من جزء واحد، إذ لا بد أن يشتملا على كلمتين على الأقل، وكل منهما يتوفر على مسند ومسند إليه، وبغيرهما لا يستقيم كلام ولا تتحقق فائدة.

---

[1] « الجملة الكبرى هي الاسمية التي خبرها جملة نحو: زيد قام أبوه، وزيد أبوه قائم، والصغرى هي المبنية على مبتدأ كالجملة المخبر عنها في المثالين ». الأنصاري (ابن هشام)، مغني اللبيب عن كتب الأعاريب، مرجع سابق، ج. 2، ص. 477.

2.2. الجملة عند الدارسين الغربيين

لقد ركز النحو التقليدي –الفرنسي خاصة والأوروبي عموما– على اعتبار الجملة تجمُّعا للكلمات من أجل تشكيل معنى تام، وهي تتميز عن القضية بحيث إن الجملة قد تتضمن عدة قضايا (في الجملة المركبة مثلا). ولكن هذا التعريف في نظر بعض الدارسين غير عملي؛ إذ يمكن أن يضمَّن المعنى في جملتين أو أكثر، كما أن المعنى قد يكون تاما دون أن يكون هناك بناء سليم للجملة (كما هي الحال مثلا بالنسبة لمتكلم بلغة غير لغته الأم)[1].

وحتى في العصر الحديث، لا يزال الاختلاف في تحديد الجملة سائدا في أوساط اللسانيين والنحاة، ولعل ذلك راجع أولا إلى صعوبة وضع تعريف جامع مانع للجملة، وإلى اختلاف مشاربهم وتوجهاتهم ثانيا؛ فمنهم من يعرفها بأنها : الوحدة الكبرى للمستوى اللساني الذي تظهر فيه العلاقات التركيبية بين العناصر، أو هي بنية حملية ثنائية تقيم علاقة بين موضوع ومحمول تتكون نواتها بالأساس من فعل متصرف، وهي تتكون من وحدات دالة في إطار تراتبي[2].

إن هذا التعريف لا يبتعد كثيرا عما ألزم به المناطقة أنفسهم من جعل الجملة عبارة عن بنية منطقية تتطابق مع العالم صدقا أو كذبا، وأساس هذه البنية هو فعل متصرف، فلا وجود للجملة إن لم يكن هناك فعل، مع العلم أن هذه الخاصية تشتهر في لغات هندوأوربية كاليونانية والإنجليزية والفرنسية، وقد لا تشتهر في لغات الأخرى كالعربية مثلا التي تكون الجملة فيها اسمية أحيانا ولا وجود للفعل فيها.

---

[1] ينظر: Dubois (Jean) et *al*., Dictionnaire de linguistique et des sciences du langage, Larousse, Paris, 1999, p. 365.

[2] ينظر: Neveu (Franck), Lexique des notions linguistiques, Nathan, Paris, 2000, p. 83

غير أن اللافت في هذا هو إشارته إلى الإطار التراتبي الذي تنتظم فيه الوحدات الدالة داخل الجملة؛ حيث إن هذا الإطار يشبه صورة قوالب: تترابط فيه الوحدات داخل القالب الأول، والوحدات داخل القالب الثاني، ثم تترابط وحدات القالب الأول بتلك التي يحتويها القالب الثاني وهكذا دواليك في شكل سلمي هرمي تراتبي.

ومن الباحثين من حاول حصر التعاريف المختلفة للجملة كما فعل **مونان Mounin** في معجمه اللساني حيث أحصى عدة تحديدات للجملة، منها احتواؤها على موضوع ومحمول المذكور آنفا، ومنها أن الجملة هي ملفوظ تام من جهة المعنى، أو هي وحدة نغمية بين وقفتين (سكوتين).

ومن جهة نظر بنوية، تعتبر الجملة قطعة من السلسلة الكلامية مستقلة من الناحية التركيبية؛ أي هي شكل لساني مستقل غير متضمَّن في شكل لساني أكبر بموجب بنية نحوية، بعبارة أخرى: هي مكوَّن لا ينضوي في مكوَّن أكثر امتدادا، وهي وحدة كبرى للوصف النحوي، لا تنتمي إلى صنف توزيعي، وبالتالي لا يمكن أن يتم الاستبدال بينها وبين مكونات أخرى[1].

أما من وجهة نظر التوليديين، فإن الجملة هي أولية قاعدية (أكسيوم) تتفرع عنها سلسلة من الرموز بواسطة قواعد تركيبية. الجملة P هي التي تتفرع من القاعدة الأولى: $P \rightarrow SV +$ SN حيث إن P هو الرمز الأولي، وSV وSN هما على التوالي: المركبان الفعلي والاسمي[2].

---

[1] ينظر: Mounin (George), Dictionnaire de linguistique, PUF, Paris, 1974, p. 262

[2] ينظر: Ruwet (Nicolas), Introduction à la grammaire générative, Plon, Paris, 1re éd., 1968, p. 47.

وقد بين **مونان Mounin** بأن الباحثين عادة ما يتجوزون في إطلاق الجملة على الملفوظ عموما، وإن كان من الأفيد حسب نظره أن تخصص الجملة لوحدات اللغة (اللسان)، وأن يخصص الملفوظ لوحدات الكلام أو أمثلته.

إن هذا التفصيل قد يتلاءم مع تعريف الجملة بأنها هي كيان مجرد يتوافق مع مخطط يكمن في ملكة المتكلم، وهي في انتظار التحيين داخل الخطاب بواسطة فعل التلفظ[1]. بالإضافة إلى أن هذا التعريف يعطي للجملة مكانتها المعرفية والاتصالية حين يبين موقعها من الملكة اللغوية، التي هي إحدى المكونات العامة للمعرفة، ويعتبرها مخططا مدخرا في الذاكرة وهو بحاجة إلى الاستحضار والاسترجاع والتجسد داخل الخطاب أثناء عملية التواصل.

وعليه، فإن الجملة من المنظور المعرفي، تحقق وحدة متكاملة مفهوميا ولسانيا على حد سواء؛ فهي تعبر، على المستوى المفهومي، عن واقعة تامة تشمل الحدث أو الكينونة كما يراها المتكلم. وعلى المستوى اللساني، تسمي الجملة فاعلا حقيقيا واحدا على الأقل، كما تذكر الحدث أو الكينونة التي يتصف بها. وبواسطة وحدات لسانية، تشير الجملة أيضا إلى الطريقة التي يرتبط بها الحدث أو الكينونة بحال المتكلم في المكان والزمان.

في التصور المعرفي أيضا، تتكون الجملة من وحدات حاملة للمعنى مرتبطة فيما بينها بطريقة نظامية. إن الجملة تصف واقعة، وتشكل الوحدة المركزية للنحو بمعناه الخاص أي التركيب. وهي تتحدد على المستوى الخطي (الكتابي) بواسطة نقطة أو علامة ترقيم أخرى، وعلى المستوى اللفظي تتميز بإطار تنغيمي خاص. إن الجملة إذن هي بناء معقد مؤلف من

---

[1] ينظر: Neveu (Franck), Lexique des notions linguistiques, op.cit, p. 83

العناصر الآتية: شبكة واقعية، مخطط بنائي مكون من المحمول وموضوعاته، عناصر خلفية توقِّع الكلام وتحدده في الإطارين المكاني والزماني[1].

إن استخدام الجملة كوحدة نظرية ووصفية عند الباحثين على رغم اختلاف مشاربهم ووجهات نظرهم يقوم على افتراضات أساسية تتكرر في سمات اللغات المختلفة وفي العديد من النماذج المنبثقة عنها، وهذه الافتراضات الأساسية هي:

- إن للجملة معنى، وغالبا ما تعرف الجملة بوحدات معنوية كاملة، أو أنها تعبر عن فكرة كاملة، وقد يفهم هذا المعنى من الجملة وحدها دون الاستعانة بمعلومات أكثر حول السياق أو المتحدث أو غيره.
- إن الجملة تعبر عن معنى جوهري موجود في اللغات كافة، فعلى سبيل المثال: الفاعل هو القائم بالحدث، والمفعول هو الذي يتأثر بالحدث، والفعل هو الذي يعبر عن الحدث، والأحداث مترابطة فيما بينها من حيث المكان والزمان، أي إن الجملة قاسم مشترك بين جميع اللغات البشرية.
- تتشكل الجملة في أية لغة من اللغات وفقا لقواعد يعرفها ويكتسبها المتحدثون الأصليون بهذه اللغات.
- تتشكل الجملة من وحدات أصغر منها كالمركبات وأشباه الجمل، وهي في الوقت ذاته أعلى مستويات التركيب القواعدي، غير أن النصوص والسياقات تشكل مستويات أعلى من الجمل.

---

[1] ينظر: Delbecque (Nicole), Linguistique cognitive, De Boeck, Bruxelles, 2e éd. 2006, pp. 105-106.

- توجد الجملة على مستويي الحديث والكتابة[1].

3. المكونات الخطية للجملة

ترتبط الجملة في بنائها بفكرة العامل ارتباطا وثيقا؛ فهو الذي يقوم بوصل مكوناتها ويجعلها نسيجا متلاحما، ويشكل معها وحدة معنوية وبناء متكاملا. وقد ركز النحاة كثيرا على فكرة العامل، وبينوا أهميتها في النحو العربي، ويذكر **الأسدي** بأن سيبويه يتناولها باستمرار، وهو يقصد بها إمكانية توليد المجالات أو إفراغها أو إنشائها لتشغلها المقولات الاسمية وترتبط بالعامل، فتدل على الوظائف النحوية. أما الجملة عنده فهي الهيأة الحاصلة من العمل، وهو الهيكل الذي يشمل مختلف الوظائف النحوية[2].

وعلى اعتبار أن الجملة قد تكون مبنية على اسم أو فعل، فقد قسمها النحاة إلى جملة اسمية وجملة فعلية؛ حيث يكون الفعل في الجملة الفعلية العنصر الفعال والمحرك في تكوين المجالات، وتتشبه به بقية العوامل كاسم الفاعل واسم المفعول في هذه المقدرة. أما في الجملة الاسمية فإن المبتدأ حسب سيبويه هو الذي يقوم بذلك الدور[3]، لذلك سنشير إلى توليد تلك المجالات وتكوينها في الجملة الفعلية ثم في الجملة الاسمية.

1.3. الكلم المولدة في الجملة الفعلية

يتمتع الفعل بين سائر أنواع الكلم الأخرى بقدرته على أن يكوِّن جملة، سواء أكانت هذه الجملة أساسية إذا ابتدئ به الكلام أم كانت فرعية إذا كانت متضمنة داخل جملة كبرى، ويرجع

---

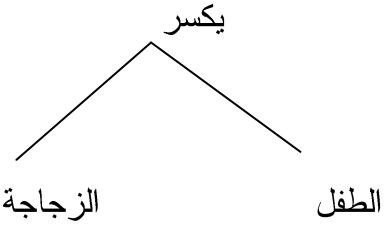
[1] ينظر: ماكارثي (ميشيل)، قضايا في علم اللغة التطبيقي، تر. عبد الجواد توفيق محمود، المجلس الأعلى للثقافة، القاهرة، ط. 1، 2005، ص. 78-79.

[2] ينظر: الأسدي (حسن عبد الغني جواد)، مفهوم الجملة عند سيبويه، دار الكتب العلمية، بيروت، ط. 1، 2007، ص. 155.

[3] يقول سيبويه: «... المبتدأ يعمل فيما بعده كعمل الفعل فيما يكون بعده. » الكتاب، مرجع سابق، ج. 2، ص. 87.

النحاة مقدرة الفعل على العمل إلى أنه يدل على الحدث، لذلك تحمل عليه بعض الأسماء المشتقة في القيام بهذا الدور كاسمي الفاعل والمفعول.

ويعتبر **تنيير Tesnière** من الباحثين اللسانيين الذين أعطوا للفعل أهمية كبيرة داخل الجملة، وجعلوه النواة التي تتعلق بها سائر مكوناتها، وقد صاغ هذه الفكرة في نظرية لسانية تسمى بنظرة التبعية أو التعلق حيث يحل الفعل في أعلى ركن أو مركب تنحدر منه بقية الأركان وترتبط به[1]، ولقد صارت هذه النظرية معتمد الدراسات اللسانية الصورية والحاسوبية في بناء الأنحاء المولدة للجمل، وفي تزويد الحاسوب بالبيانات والعمليات اللازمة في ذلك.



ونظرا إلى أن الحدث يصدر عن فاعل، ويقع على موضوع معين بطريقة ودرجة محددتين وانطلاقا من أسباب خاصة، ويتم في أحياز زمانية ومكانية، تتعلق بالفعل مجموعة من المكونات هي الفاعل، والمفاعيل: المفعول المطلق، والمفعول به (ويشمل كذلك المنادى والمستثنى)، والمفعول فيه، والمفعول له، والمفعول معه، والحال، والتمييز، وهي تعبر عن كل تلك العناصر والعلاقات.

وبالتالي، يمكن تحديد عملية توليد المكونات داخل الجملة بأنها هي الوظيفة الأساسية للفعل، ويؤكد سيبويه على هذه الفكرة قائلا: « هذا باب الفاعل الذي يتعداه فعله إلى مفعولين،

[1] ينظر: ديتربونتنج (كارل)، المدخل إلى علم اللغة، تر. حسن بحيرى، مؤسسة المختار، القاهرة، ط. 1، 2003، ص. 393-194.

فإن شئت اقتصرت على المفعول الأول، وإن شئت تعدى إلى الثاني كما تعدى إلى الأول »[1]، فبين أن الفعل يولد الاسم الذي يقوم بوظيفة الفاعل، كما أنه قد يتعدى الفاعل فيولد مفعولا أو مفعولين.

واللافت هنا أن مفهوم التعدي يختلف عما اصطلح عليه النحاة المتأخرون خاصة، فليس المقصود به احتياج الفعل إلى مفعول به في مقابل اللازم الذي لا يحتاج إلى مفعول به[2]، بل هو مفهوم تكويني حيث يقوم الفعل بتكوين الجملة فيستدعي الفاعل والمفعول وبقية المكونات على حسب مقتضى الحال ومقاصد المتكلم، وتبعا لشبكة العلاقات الدلالية التي يثيرها هذا الفعل.

كما يشير سيبويه إلى قدرة الفعل على إيجاد منازل أو محال تشغلها المكونات التي يسعى الفعل إلى توليدها منها مجال الفاعل ويليه مجال المفعول ثم مجال الحال، ولا تستطيع المفردة من الناحية البنائية أن تحل في مجال قد شغلته مفردة أخرى؛ يقول سيبويه: « وذلك قولك: ضربت عبد الله قائما، وذهب زيد راكبا... فالاسم الأول المفعول في ضربت قد حال بينه وبين الفعل أن يكون فيه بمنزلته، كما حال الفاعل بينه وبين الفعل في ذهب أن يكون فاعلا. »[3] لذا، فإن الفعل كما يتميز بسمات معجمية تجعله يرتبط ببعض المكونات اللسانية دون بعض، كذلك ينفرد بسمات بنائية تدفعه إلى توليد المجالات والتحكم بالكلم الشاغلة لتلك المجالات من أجل القيام بوظائف معينة داخل الجملة.

---

[1] سيبويه (أبو بشر)، الكتاب، مرجع سابق، ج. 1، ص. 77.

[2] ينظر: ابن عقيل (بهاء الدين)، شرح ابن عقيل، ت. حنا الفاخوري، دار الجيل، بيروت، ط. 5، 7997، ج. 1، ص. 418.

[3] سيبويه (أبو بشر)، الكتاب، مرجع سابق، ج. 1، ص. 44.

وبالتالي، يتم فهم الجملة عبر مستويين: مستوى السمات البنائية حيث تظهر فيه منزلة الفعل ويليها بقية مجالات الكلم، ومستوى السمات المعجمية حيث تظهر فيه احتياجات الفعل إلى مكونات معينة وكذا عملية الملاءمة بينه وبين تلك المكونات[1].

2.3. الكلم المولدة في الجملة الاسمية

إن المبتدأ في الجملة الاسمية لا يشتمل على القدرة التي يتمتع بها الفعل في الجملة الفعلية على توليد المجالات التي تحتلها مختلف العناصر من أجل أن تؤدي وظائف معينة، وإنما تقتصر قدرته على إحداث مجال تلازمي دلالي يربطه بالخبر، رغم ذلك قد يكون المبتدأ في حالة شبيهة بالفعل، فيمتلك قدرة على إنشاء المجالات وتكوينها أحيانا.[2]

مثال ذلك اسم الإشارة الواقع مبتدأ الذي ينصب الحال نظرا لشبهه بالفعل في دلالته على التنبيه والتعريف؛ يقول سيبويه: « وأما النصب فقولك: (هذا الرجل منطلقا)، جعلت (الرجل) مبنيا على (هذا)، وجعلت الخبر حالا له قد صار فيها، فصار كقولك: (هذا عبد الله منطلقا)... فكأن ما ينتصب من أخبار المعرفة ينتصب على أنه حال مفعول فيها، لأن المبتدأ يعمل فيما بعده كعمل الفعل فيما يكون بعده، ويكون فيه معنى التنبيه والتعريف، ويحول بين الخبر والاسم المبتدأ كما يحول الفاعل بين الفعل والخبر... »[3]. لقد حاجج سيبويه على أن اسم الإشارة المبتدأ ينصب الحال لأنه يشبه الفعلين (أنبِّه) و(أعرِّف) في قدرتهما على نصب الحال، بالإضافة إلى

---

[1] ينظر: الأسدي (حسن عبد الغني جواد)، مفهوم الجملة عند سيبويه، مرجع سابق، ص. 162.

[2] ينظر: نفسه، ص. 775.

[3] سيبويه (أبو بشر)، الكتاب، مرجع سابق، ج. 2، ص. 87.

وجود شبه شكلي بينهما إذ يفصل بين كل منهما عنصر آخر: الفاعل في الجملة الفعلية والخبر في الجملة الاسمية.

4. اللسان العربي

أتعرض في هذا البحث إلى دراسة لسانية حاسوبية لمسار توليد الجمل في اللسان العربي، لذلك فمن الضروري تعريف هذا المفهوم الأخير، خاصة أن هذا البحث ليس بحثا نظريا يرمي إلى دراسة توليد الجمل بصفة عامة وفي جميع الألسنة، ولكنه بحث تطبيقي أسعى من خلاله إلى استثمار المنجزات الحديثة في هذا المجال لدراسة اللسان العربي.

وبالتالي، فاللسان العربي هو اللسان الذي تنطلق منه الدراسة، وهو مفهوم جامع ينطوي على مجموعة اللغات والتنوعات والاستعمالات العربية والمختلفة أحيانا، والتي يُتكلم بها في العالم العربي وخارجه. أحاول أن تطرق إلى الخريطة الجغرافية للمتكلمين بالعربية، ثم أعرج على تنوعاتها.

1.4. الخريطة الجغرافية

يتكلم بالعربية اليوم أكثر من 422 مليون شخص في 22 دولة على الأقل:

- الجزيرة العربية: السعودية، البحرين، الإمارات المتحدة، عمان، قطر، اليمن؛
- الشرق الأوسط: العراق، الأردن، الكويت، لبنان، فلسطين، سوريا؛
- إفريقيا: الصومال، جيبوتي، جزر القمر، مصر، السودان، ليبيا، تونس، الجزائر، المغرب، موريتانيا.

كما أنها اللغة التي يرجع إليها أكثر من مليار مسلم عبر العالم. اللغة العربية هي اللغة التي يتكلم بها العرب، وهي لغة سامية (مثل العبرية والآرامية والسريانية)، نشأت في جزيرة العرب ثم انتشرت وتوسعت بفضل الفتوحات الإسلامية[1].

يتضمن مصطلح "اللسان العربي" عدة تنوعات (لغات[2] ولهجات واستعمالات) كلها مشتركة من حيث النظام اللساني العام، ومتقاربة من حيث المعجم اللغوي منها:

– العربية الفصيحة: وهي تشمل لغات العرب الفصحاء والتي وردت في القرآن والشعر الجاهلي، وتلحق بها العربية القياسية الحديثة التي تشكل اللغة المكتوبة في الدول العربية، واللغة الرسمية والمشتركة بينهم حيث تستخدم في الإدارة والتعليم والإعلام والعلوم والفنون بالإضافة إلى الأدب.

– العربية اللهجية: وهي مجموعة من اللهجات والعاميات التي يتكلم بها في الحياة اليومية في العالم العربي وخارجه، والمتفاوتة بين الدول العربية وحتى داخل القطر الواحد، لكن يمكن أن يُجمع المتقارب منها في مجموعات كبرى أشهرها اللهجة المغاربية التي تضم اللهجات المستعملة في المغرب الكبير.

---

[1] ينظر: مقال اللغة العربية في موسوعة ويكيبيديا الحرة على العنوان: http//ar.wikipedia.org/wiki/اللغة_العربية (تاريخ الاطلاع على الصفحة: 2033/07/20).

[2] المقصود باللغات هنا على حد تعبير الخليل: اختلاف الكلام والمعنى واحد، وعليه كانت تطلق على اللهجات العربية القديمة: لغات العرب. لذلك آثرت هنا استخدام مصطلح "اللسان" الذي يحيل على التوحد كما في قوله تعالى: "وَهَذَا لِسَانٌ عَرَبِيٌّ مُبِينٌ"، في مقابل اللغة التي تتعدد عناصر مدلولاتها (لغة المكتوب/لغة المنطوق، لغة الشارع، لغة كاتب/كاتب آخر، لغة الإعلام...). وللحاج صالح رأي في التفريق بين اللهجة واللغة يخالف فيه الباحثين المحدثين إذ يعتبر اللهجة تنوعا إقليميا قد يمس النظام أحيانا، في حين أن اللغة هي تنوع في بعض جوانب الأداء والاستعمال (كالفتح والإمالة) ولا يصل إلى درجة اللهجة.

2.4. النحو الصوري للغة العربية

من الضروري بالنسبة للدراسات اللسانية المعمقة ولتلك الدراسات الخاصة بالمعالجة الآلية، بلورة نحو صوري للغة العربية. غير أن هناك عدة مشاكل تصادف الباحثين في هذه العملية منها قلة الموارد والمدونات والذخائر اللغوية المخزنة في شكلها الرقمي والتي يعتمدون عليها في دراستهم، وكذا الإطار الصوري الذي ينبغي اتباعه لإنجاز تلك العملية.

- **موارد اللغة العربية**

في انتظار أن ترى النور مختلف المشاريع العربية (ومنها مشروع الذخيرة العربية) التي تعمل على تدوين النصوص والمصادر العربية في الحاسوب، يقيم الباحثون دراساتهم على ما توافر لديهم من مصادر للغة العربية. بالمقابل، تتولى بعض الجامعات الغربية إقامة مشاريع لتطوير نحو صوري للغة العربية في إطار لسانيات المدونة[1] منها:

- **مشروع الذخيرة العربية ببنسلفانيا *Penn Arabic Treebank***: وهو مشروع تتبناه جامعة بنسلفانيا الأمريكية من أجل تطوير نحو خاص بالمعالجة الآلية للغة العربية، وهو يتضمن في طياته تحشية لأقسام الكلام والعلاقات بين المحمول والفاعل في شكل نحو مركبي[2].

---

[1] **لسانيات المدونة *linguistique de corpus*** هي جمع البيانات اللسانية في مدونات وذخائر من أجل الاستفادة منها في مختلف مجالات البحث، ومع ظهور الحاسوب، عمل الباحثون على تزويده بمختلف المعطيات التي تحتوي عليها اللغة. ينظر:
Malmkjær (Kirsten), The linguistics encyclopedia, Routledge, London, 1991, 1st ed., p. 73.

[2] ينظر: Ann Bies and Mohamed Maamouri, PENN ARABIC TREEBANK GUIDELINES, 2003, http//www.ircs.upenn.edu/arabic/Jan03release/guidelines-TB-1-28-03.pdf
(تاريخ الاطلاع على الصفحة: 2033/03/21).

إن النحو المركبي المعتمد هو مجموعة من القواعد من الشكل: X → Y، والتي تخضع لضوابط منها:

- هي قواعد إعادة كتابة الرموز الموجودة على يسار السهم في الرموز الموجودة على يمينه؛
- هي قواعد توسع وامتداد نظرا لأنها تعوض عنصرا وحيدا بعنصر أو عدة عناصر[1].

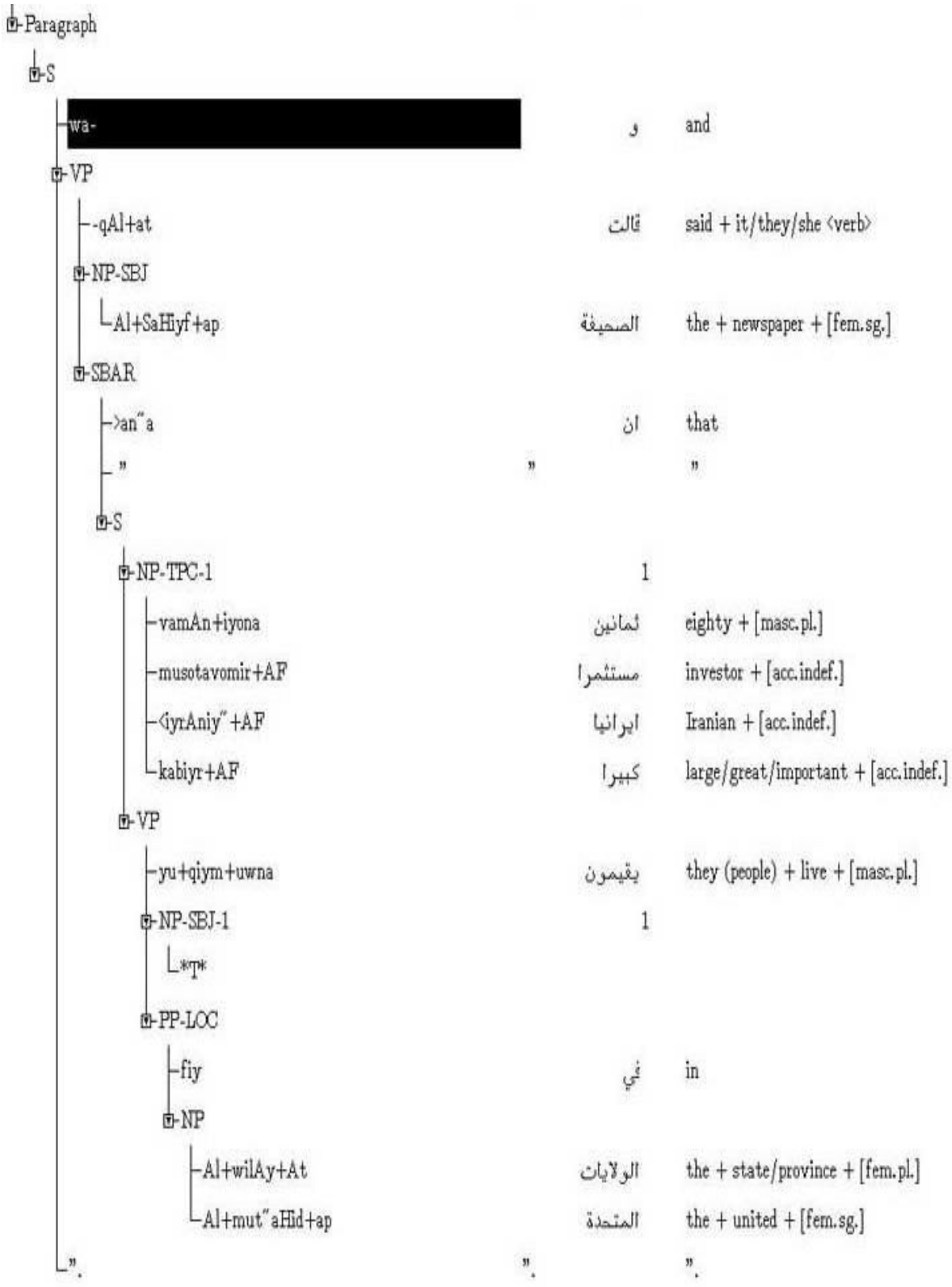
- **مشروع الذخيرة العربية ببراغ*Prague Arabic Dependency Treebank*** : وهو مشروع تتبناه جامعة براغ التشيكية، وهو يعتمد على نحو التعلق[2]، ويهدف إلى تحويل المشجرات التركيبية للمشروع الآنف الذكر إلى مشجرات تعلق بطريقة آلية[3].

---

[1] ينظر: Bach (Emmon), Introduction aux grammaires transformationnelles, trad. Robert Sctrick, Armand Colin, Paris, 1973, pp. 53-54.

[2] يرجع هذا النحو في أصله إلى بحوث **تنيير Tesnière** الذي حاول الكشف عن الحدث البنوي العميق الذي يلوح خلف المظهر الخطي للغة في السلسلة الكلامية حيث يجعل الكلمات التي تكوّن الجملة ضمن مقولات، ويحدد صلات التعلق التي تربط بين تلك الكلمات، وتمثل هذه الصلات أو العلاقات بواسطة رسم بياني يدعى **الستيما *stemma***. ينظر: Tesnières (Luc)))) sssssssse ddesss œexxxxxccccccaccaaeeeeeeeeeeeePeeP444444444444444

[3] ينظر: Hajic (Jan) et *al*., Prague Arabic Dependency Treebank: Development in Data and Tools, http://ufal.mff.cuni.cz/padt/PADT_1.0/docs/index.html.
(تاريخ الاطلاع على الصفحة: 2033/03/21).



الشكل2: نموذج من مشجر تعلقي

هذا الشكل هو مخطط لجملة "وقالت الصحيفة أن ثمانين مستثمرا إيرانيا كبيرا يقيمون في الولايات المتحدة"؛ يحيل الرمز S إلى الجملة وهو مساو لـــ VP (مركب فعلي) إذا لم يتقدم على الفعل شيء، الفاعل هو NP-SBJ (وNP-TPC إذا تقدم على الفعل). تبدو في هذا المخطط تحشية مختلف المقولات المؤلفة للجملة بواسطة مداخل معجمية، وهو يبين العلاقات بين المحمول والفاعل VP-NP في شكل تراتبي يربط المركب بمكوناته.

مثال: الولايات المتحدة (NP) --- في (PP) ---- يقيمون (VP).

- **الإطار النظري للنحو الصوري العربي**

يقدم النحو العربي في مختلف المصادر منذ سيبويه على أنه نحو تعلقي، وهناك كثير من الدراسات الحديثة التي حاولت أن تطور نظريات سيبويه والنحاة العرب، وأن تنحوا منحى صوريا كأعمال الحاج صالح، وانصبت بالمقابل جهود كثير من اللسانيين العرب في دراسة الجملة في إطار نحو **بنية المركبات *Phrase-Structure Grammar***، كما يظهر ذلك عند الفاسي الفهري الذي يحاول بلورة وصف تأويلي لبنية الجملة في اللغة العربية، حيث لا يعتمد فقط على ما يلاحظ من الظواهر المتعلقة برتب المكونات والمركبات في الجملة، بل حتى على تفسير تلك الظواهر غير الملاحظة من أجل وضع النظام النحوي الذي يضبطها[1].

من ناحية أخرى، ينطلق **ديترس Ditters** في تحليله الآلي للغة العربية من مقاربة إثباتية؛ لا يعتمد فيها على التأويل، ولكن على وصف البنى التركيبية والدلالية. لذلك، يحتاج

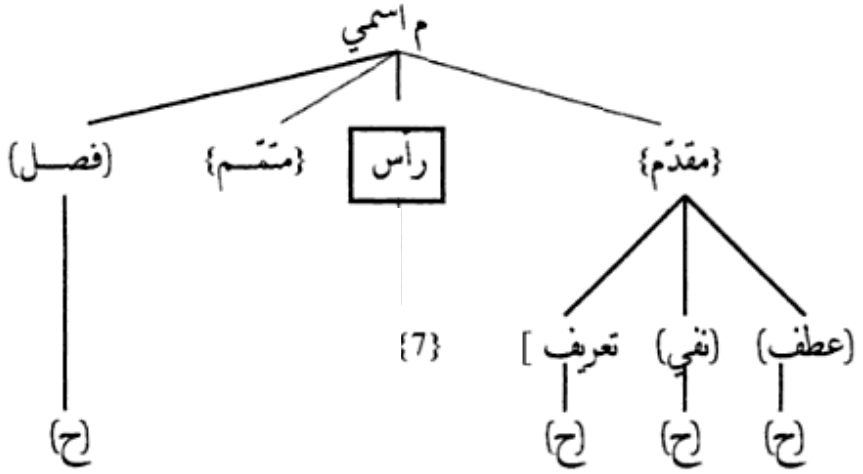
[1] ينظر: الفاسي الفهري (عبد القادر)، اللسانيات واللغة العربية، دار توبقال، المغرب، ط. 1، 1985، ص. 303.

حتما إلى مدونات واسعة لإتمام وصفه، هذا الوصف الذي يبلور عن طريق وظائف خاصة وعلى شكل مكونات مباشرة. وعليه، يفرق هذا الباحث بين نوعين من الجمل:

- جملة اسمية تكون مبينة وتتضمن وظيفة محورية[1] وتكون مقولتها الأساسية مركبا اسميا، أو تكون شارحة ومقولتها الأساسية هذه المرة إما مركب اسمي أو نعت أو حال أو جار ومجرور أو جملة.
- جملة فعلية مع وظيفة حملية يتمثل عنصرها الأساسي في المركب الفعلي، والمكونات الأخرى تعرّف بواسطة البنية التعلقية للفعل.

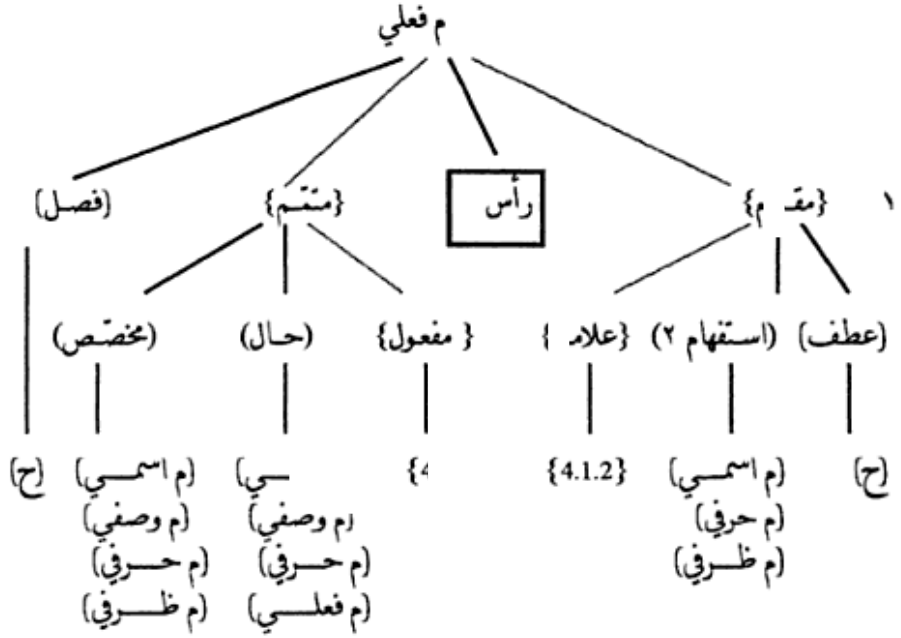
تتكون الجمل من مركبين أساسيين هما المركب الاسمي والمركب الفعلي، أما المركب الاسمي (كما يلاحظ من خلال الشكل) يتكون من عدة وظائف؛ فبالإضافة إلى الرأس، هناك مجموعة من العناصر ترتبط به وتؤدي وظائف محددة هي: المقدم الذي يتفرع إلى ثلاثة أدوار: العطف والنفي والتعريف وغالبا ما تؤدى هذه الأدوار بالحروف، ثم المتمم وهي وظيفة تؤدى بالمركبات الوصفية أو غيرها، ووظيفة الفصل الخاصة بعلامات الترقيم أو ووصل المركب الاسمي بعناصر أخرى والتي تؤدى بالحرف.

[1] « المحور هو المكون الدال على المحدث عنه داخل الحمل. » المتوكل (أحمد)، دار الثقافة، الدار البيضاء، ط.1، 1985، ص. 69.



الشكل3: بنية المركب الاسمي

وأما المركب الفعلي فهو المقولة القادرة على تحقيق وظيفة الإسناد في الجملة الفعلية. الوظيفة المركزية في المركب الفعلي هي الرأس المتحقق في صورة الفعل. في مقدم هذا المركب قد نجد وظائف ثانوية مثل العطف والاستفهام وعلامة الزمن (الماضي مثلا) أو صيغة الفعل (الأمر مثلا). بعد الرأس يمكن أن نصادف وظائف ثانوية متممة للفعل (كالفاعل والمفعول)، والحال ومخصصات أخرى، أما الفصل فهو الموقع الخاص بعلامات الترقيم أو مدخل لرابط يصل المركب الفعلي بغيره من العناصر أو المركبات. يمثل **ديترس Ditters** لبنية المركب الفعلي كالآتي:



الشكل4: بنية المركب الفعلي

كما نرى من خلال هذه البنية، قد تؤدى أدوار الاستفهام والحال والمخصص بعدة مركبات؛ وظيفة الاستفهام تقوم بها مركبات اسمية (أي) وحرفية (هل) وظرفية (أين)، ووظيفة الحال تقوم بها مركبات اسمية ووصفية (كاتبا) وحرفية (بالسيارة) وفعلية، ووظيفة المخصص تقوم بها مركبات اسمية ووصفية وحرفية وظرفية، أما وظيفتا العطف والفصل فغالبا ما تؤديان بالحرف[1].

في نهاية هذا الفصل الذي أردته مدخلا تعريفيا بأبرز المفاهيم التي يتناولها البحث خاصة مفهوم الجملة، هذا المفهوم الذي تعرضت له اتجاهات متباينة ضمن مناهج متعددة، أختار من بينها النموذج الذي تبناه الحاج صالح في وصفه للجملة العربية ضمن المنهج الصوري الذي

---

[1] ينظر: f f Syxxxx e Soooo◉oorr a Deccooooooff ooooo ‘‘TTee EEeeaaar)) tttt e f ccc cc ’’’’*in* Langues et Littératures du Monde A’’’’ ’’ e æeesssss sss ssnnsnnnnn et littératures du monde arabe, 2001, pp. 115-151.

يجنح نحو تحليل الجملة إلى مكوناتها المباشرة بالاعتماد على قواعد إعادة الكتابة لأقيم عليه تطبيقا حاسوبيا لتوليد الجمل، لكن قبل ذلك أرى من اللازم التطرق إلى الدراسة اللسانية الحاسوبية قبل الشروع في التطبيق المذكور.

الفصل الثاني:

الدراسة اللسانية الحاسوبية

الفصل الثاني: الدراسة اللسانية الحاسوبية

إن الدراسة اللسانية الحاسوبية تعالج اللغة حسب تصورات ومناهج خاصة ليست بالضرورة مختلفة عن التصورات والمناهج اللسانية التقليدية، ولكن أهداف تلك الدراسة هي التي تعطيها طابعا مميزا؛ حيث إنها تحاول إقامة أوصاف صورية صارمة لمختلف الظواهر اللغوية من أجل تزويد الآلة بشتى المعارف والعمليات الموجودة في اللغة.

أحاول في هذا الفصل أن أتعرض إلى تعريف اللسانيات الحاسوبية، وبيان محاورها ومناهجها، وأتخير منها ما يكون سندا لي في دراستي لتوليد الجمل في اللسان العربي.

1. اللسانيات الحاسوبية: المفهوم والمنهج

لقد كانت اللسانيات الحاسوبية في بداياتها تعتمد التحليل الإحصائي للمفردات اللغوية في كتاب معين أو لدى كاتب معين لإعداد فهارس أبجدية لتلك المفردات وتحديد تواترها في مؤلفاته، ثم خطت خطوة مهمة جداً في اختصار الزمن عندما أعدّت معاجم إلكترونية أحادية اللغة أو ثنائيتها أو متعددة اللغات.

وهكذا تم خلال الستينات من القرن العشرين وما تلاها حتى الآن إنتاج برامج وأنظمة للترجمة الآلية وغيرها من التطبيقات اللسانية الحاسوبية، كان بعضها تجارياً بحتاً، تُغذّى به الحواسيب لتترجم جملا مكتوبة أو منطوقة ومصطلحات كاملة في مجالات محددة أهمها السياحية والتجارية والمرتبطة بالخدمات كمصطلحات التحية والاستفسار عن الأسعار والأماكن

والزمن وغيرها[1]. لكن كغيرها من التخصصات الحديثة والأجنبية المنشأ، تتعدد المصطلحات الدالة على اللسانيات الحاسوبية والتي عكف الباحثون على استخدامها سواء أكان ذلك في المراجع الأجنبية أم العربية، مما يسبب مشكلا في توظيف المصطلح وترجمته إلى اللغة العربية.

1.1. مشكل المصطلح والترجمة

في أول مؤتمر دولي يقام بشأنها سنة 1955، عُرِّفت اللسانيات الحاسوبية بأنها: «علم جديد تتقاطع فيه اللسانيات مع جهاز صوري تفرزه العلوم المنطقية الرياضية ويخضع للقيود التي تفرضها الآلات المعدة للمعالجة الآلية للمعلومة، ويؤدي البحث في هذا المجال إلى إنشاء نموذج خوارزمي»[2]. لقد حددت في هذا التعريف ملامح اللسانيات الحاسوبية التي تتمثل في تلاقي جهازين مفهوميين حديثين نوعا ما هما: اللسانيات في تطورها المتزايد، والعلوم المنطقية الرياضية في رؤيتها الصورية، بمجال تقني حديث النشأة كذلك هو المعالجة الآلية للمعلومة. كما أعلن هذا التعريف عن الهدف من ظهورها وهو صياغة نموذج خوارزمي وهو خطاطة منطقية رياضية للغة قابلة لأن تشغل في الحاسوب.

---

[1] ينظر: الواسطي (سلمان داود)، "التفاعل بين الإنسان والآلة في الترجمة الحاسوبية"، مجلة التعريب، المركز العربي للتعريب والترجمة، دمشق، ديسمبر 2000، العدد 20، ص. 7.

[2] ينظر: Léon (Jacqueline), "De la traduction automatique à la linguistique computationnelle. Contribution à une chronologie des années 1959-1965", Traitement Automatique des Langues N° spécial trentenaire, 1992, vol.33, n° 1-2 :25-44.

والخوارزم مجموعة محددة من خطوات منطقية وحسابية تحدد المنهج لحل مشكلة معينة، وهذا الاسم مشتق من اسم محمد بن موسى الخوارزمي (ق. 8هـ) الذي أسس المنهج الرياضي لحل المسائل. ينظر: معجم الحاسبات، مطبوعات مجمع اللغة العربية، القاهرة، ط.2، 1995، ص. 7.

وفي الأدبيات العربية، يحيل هذا المصطلح (اللسانيات الحاسوبية[1]) عادة إلى المجال الذي ترتبط فيه اللسانيات أو علوم اللغة بعلوم الحاسوب، كما يظهر من كلام أحد الباحثين: «مما لا شك فيه أن معالجة اللغة العربية كلغة من اللغات الطبيعية تدخل في علم مخصوص وليد التطورات التكنولوجية المتقدمة ألا وهو اللسانيات الحاسوبية، مجالها البحثي دقيق وجديد يعرض لآخر النظريات والتطبيقات الحاسوبية المجربة على جميع اللغات الطبيعية»[2]، ويشدد على ارتباط هذا التخصص بالتكنولوجيا والإعلام الآلي فيقول: « يلتقي فيه الجانب النظري اللساني بكل خلفياته المعرفية والمنهجية والجانب التقني المعلومياتي بكل تطوراته ليصوغ ما اصطلح عليه بالهندسة اللسانية أو تكنولوجيا اللسان »[3] إلا أننا نلاحظ إيراده لمصطلحين آخرين وجعلهما مرادفين له وهما: الهندسة اللسانية وتكنولوجيا اللسان، مما يدل على أن هذا الباحث قد أعطى للسانيات الحاسوبية طابعا تقنيا شديد الارتباط بالآلة، وتؤكّد على ذلك طريقة صياغة هذا المصطلح (اللسانيات الحاسوبية)؛ فقد تم وصفه بـــ"الحاسوبية" التي تشير إلى نسبته وتعلقه بالحاسوب وهي الآلة التي تتجلى فيها معالجة المعلومات بطريقة آلية.

وينحو **عبد الرحمن الحاج صالح** هذا المنحى في دلالة مصطلح "اللسانيات الحاسوبية" على الحقل الذي تمتزج فيه اللسانيات بالمعلوميات حيث يقول: « إن الدراسات والبحوث العلمية في اللسانيات الرتابية (الحاسوبية) ازدهرت في الوطن العربي في هذه الآونة وتكاثر إلى حد ما

---

[1] يستخدم الكثير من الباحثين العرب "اللسانيات الحاسوبية" مقابلا للمصطلح الفرنسي: *linguistique informatique*، وكذا للمصطلح الإنجليزي: *computational linguistics*.

[2] غازي (عز الدين)، "اللسانيات الحاسوبية واللغة العربية"، الحوار المتمدن – العدد: 3639 – 2006/08/11، مجلة إلكترونية على العنوان: www.alhewar.org. (تاريخ التصفح: 2011/07/15).

[3] المرجع نفسه.

الباحثون في هذا الميدان الذي تتلاقى فيه علوم الحاسوب وعلوم اللسان، وهو ميدان علمي وتطبيقي واسع جدا كما هو معروف إذ يشمل التطبيقات الكثيرة كالترجمة الآلية، والإصلاح الآلي للأخطاء المطبعية، وتعليم اللغات بالحاسوب... »[1]. غير أن هذه التطبيقات الحاسوبية الكثيرة التي تعالج اللغة العربية ليس من اليسير أن تجمع في أصول واحدة، وأسسها الإبستمولوجية غير واضحة، وبالتالي لم توضع لها المقدمات التعليمية التي تسهل على القارئ العربي المتعلم أو الباحث أن يستفيد منها.

رغم ذلك، يُستنتج مما تقدم أن اللسانيات الحاسوبية هي مجال تتداخل فيه التصورات اللسانية والحاسوبية وتتلاقح لتشكل نظريات تعمل على معالجة الوقائع اللغوية وفق منهج حاسوبي لتتمخض عن ذلك تطبيقات متعددة تشمل تلك الوقائع اللغوية لكن في إطارها الآلي. وبالتالي، وإن كانت اللسانيات علما متجذرا في الفكر الإنساني، غير أن ارتباطها بالحاسوب هو من ابتداع القرن العشرين عصر ثورة المعلومات.

ويرجع السبب في ذلك الارتباط إلى كون الحواسيب تمثل أوج ما بلغه التقدم التكنولوجي، وأهم ما تحتاج إليه الحياة المعاصرة لأنها تساعد على حل كثير من مشكلاتها المعقدة. ويتم ذلك بالتواصل مع الحواسيب عبر لغة خاصة استفاد الباحثون من دراسة اللغات الإنسانية في تطويرها، وعلم اللسانيات الحاسوبية ناشئ من هذا التواصل مع الحواسيب، ويخدم الأهداف المتعلقة به والتي تنحصر في حل المشكلات المعقدة التي تتصل بحوسبة اللغة[2].

---

[1] الحاج صالح (عبد الرحمن)، بحوث ودراسات في اللسانيات العربية، مرجع سابق، ج. 1، ص. 230.

[2] ينظر: استيتية (سمير شريف)، اللسانيات: المجال والوظيفة والمنهج، عالم الكتب الحديث، الأردن، ط. 1، 2005، ص. 527.

ويميل إلى ذلك **نهاد الموسى** مع تفصيل أكثر حيث يعتبر اللسانيات الحاسوبية نظاما بينيا بين اللسانيات وعلم الحاسوب يعنى بحوسبة جوانب الملكة اللغوية[1]. وهي تتألف من مكونين أحدهما تطبيقي والآخر نظري، فأما المكون التطبيقي من اللسانيات الحاسوبية فيهتم بالناتج العملي لنمذجة الاستعمال الإنساني للغة، وهو يهدف إلى إنتاج برامج مزودة بكثير من المعارف اللغوية، وتمثل الترجمة الآلية إحدى غاياته الأساسية. وأما المكون النظري من اللسانيات الحاسوبية فيتناول قضايا في اللسانيات النظرية أبرزها إقامة نظريات صورية للمعرفة اللغوية التي يحتاج إليها الإنسان لتوليد اللغة وفهمها، كما تُطور اللسانيات الحاسوبية نماذج صورية تستجمع وجوه الملكة اللغوية الإنسانية وتترجمها إلى برامج حاسوبية[2].

مما سبق، يبدو أن الباحثين العرب قد ترجموا هذا المصطلح من المصطلح الإنجليزي ***computational linguistics***، الذي يشير إلى ذلك المجال الذي يتفرع من اللسانيات وعلوم الحاسوب، ويشمل التفاعلات بين اللغة الإنسانية والحواسيب، وهو يتضمن تحليل النص المكتوب والخطاب الشفوي، ترجمة نص أو منطوق من لغة لأخرى، استعمال اللغات الإنسانية (لا الحاسوبية) للتواصل بين الحواسيب ومستخدميها، كما أنه يتضمن نمذجة[3] النظريات اللسانية واختبارها[4].

---

[1] ينظر: الموسى (نهاد)، «حصاد القرن في اللسانيات» ضمن: حصاد القرن: المنجزات العلمية والإنسانية في القرن العشرين، مؤسسة عبد الحميد شومان، الأردن، ج. 2، ص. 47.

[2] ينظر: الموسى (نهاد)، العربية: نحو توصيف جديد في ضوء اللسانيات الحاسوبية، المؤسسة العربية للدراسات والنشر، بيروت، ط. 1، 2000، ص. 54.

[3] النمذجة هي عملية هدفها **كشفي *heuristique***، تتضمن إنشاء النماذج، وذلك بنقل المعطيات وملاحظتها ومن ثم وصف مختلف السيرورات من خلال لغة مناسبة وصورية. ينظر: Dufaye (Lionel), Théorie des opérations énonciatives et modélisation, éd. Ophrys, 2009, p. 38.

[4] ينظر: Fromkin (V.) and *al.*, An introduction to language, Wadsworth, 7th ed., USA, p. 412

وقد يجعله بعض الباحثين مرادفا للمعالجة الآلية للغة الطبيعية؛ إذ أن أبرز مهمة للسانيات الحاسوبية هو إنشاء برامج حاسوبية من أجل معالجة الكلمات والنصوص في اللغة الطبيعية، وهذه المهمة تتلاقى مع مهمة مجال المعالجة الآلية للغة الطبيعية. غير أنه لا يخفى بأن عملية إنشاء البرامج الحاسوبية المعدة لمعالجة الكلمات والنصوص، خاصة في الأبحاث الأولى، لم تكن تعتمد في أغلب الأحيان على ما تقدمه اللسانيات من معلومات حول الوقائع اللغوية.

لكن الظاهر أن منحى اللسانيات الحاسوبية هو لساني أكثر منه حاسوبي بمعنى أن الباحثين فيها يهتمون بالوصف الصوري للغة بدلا من اهتمامهم بالمشاكل الخوارزمية التي يمكن أن تُصادف لدى القيام بعملية الصورنة، لذلك فإن هذه المشاكل موكولة إلى المعلوميات حتى تحلها. بالإضافة إلى ذلك، فإن الخوارزميات والبرمجيات المتصلة بها والتكنولوجيات البرمجية يمكن أن تتعدد بحسب الأطر النظرية الحاسوبية وكذا بحسب الجوانب التقنية للآلة، في حين أن المبادئ اللسانية الأساسية ومناهجها الوصفية هي أكثر ثباتا[1].

بالمقابل، فإن المصطلح الفرنسي المستعمل للدلالة على ذلك المجال الذي تتداخل فيه علوم اللغة وعلوم الحاسوب هو ***linguistique informatique*** والذي تمكن ترجمته –في رأيي– إلى اللسانيات المعلومية إلى جانب اللسانيات الحاسوبية، خاصة إذا اعتبرنا بأن

---

[1] ينظر: Bolshakov (I.) And Gelbukh (A.), Computational linguistics: models, resources, applications, Instituto Politecnico Nacional, Mexico, 2004, p. 25.

المصطلح العربي المقابل لـــ***informatique*** هو المعلوميات[1]، وهو الذي يشير إلى مجموعة التخصصات العلمية والتقنية المطبقة على معالجة المعلومة باستخدام وسائل آلية[2].

لقد بين **راستيه Rastier** أن الارتباط بين علوم اللغة وعلوم الحاسوب له طرق ثلاثة: الطريقة الأولى يكون فيها التحليل اللساني أولوية بالنسبة للمعالجة المعلومية أو الحاسوبية، ويسمح هذا النوع بتحليل أولي للمدونة تبعا للمهمة المنتظر تنفيذها من الحاسوب. أما الطريقة الثانية فيوجه فيها التحليل اللساني التحليل المعلومي في إطار استراتيجية استعمال البرامج الحاسوبية. وفي الطريقة الثالثة تقوم اللسانيات بتأويل نتائج المعالجة. في أفضل الأحوال، تتدخل اللسانيات قبل التشغيل المعلومي وأثناءه وبعده[3].

الطريقة الأولى (قبلية)	الطريقة الثانية (موازية)	الطريقة الثالثة (بعدية)
التحليل اللساني للمدونة بحسب المهمة المنتظرة من الحاسوب	توجيه التحليل اللساني للتحليل المعلومي عند البرمجة	تأويل نتائج المعالجة

الجدول4: طرق الارتباط بين علوم اللغة وعلوم الحاسوب

[1] المعلوميات على غرار الرياضيات لا المعلوماتية كما يشيع على الألسنة لما فيها من مخالفة للقياس.

[2] ينظر: aaaaa aaaaccccdiccccc e eeeoo ooooo oooooooooo8888888888888

[3] ينظر: Rastier (François) et *al.* Sé aaaalye s ss ss rr rrrrrrrr Masson, Paris, 1994, p. 2.

بالرجوع إلى موضوعها، تتناول اللسانيات المعلومية[1] بالدراسة أوصافا نسقية صورية مرتبطة بالنظريات اللسانية المبنية على التفكير في اشتغال اللغة المنطوقة عند البشر (وهي اللغة الطبيعية)، وتتم هذه الدراسة بالاعتماد على وسائل معلومية حاسوبية. وعليه، تهتم اللسانيات المعلومية، من جهة، بالمقاربات النظرية التي تأخذ بعين الاعتبار أبنية الجمل على أنها مواضيع معقدة بالأساس، مثلا على مستوى تأويلاتها الدلالية، ومن جهة أخرى، تهتم بترجمة تلك النظريات إلى لغة اصطناعية عبارة عن صورانيات أو أمثلات لسانية تسمح بتشغيلها معلوميا في الحاسوب[2].

وبالتالي، يتضمن موضوع اللسانيات المعلومية تلك الأوصاف اللسانية التي تتسم بالنسقية حيث تشكل عناصرها من أصوات وكلمات وجمل وغيرها وما يقابلها من مدلولات نظامية واطرادا في ظهورها كوقائع لغوية، كما تتسم بالصورية حيث تُستخرج من تلك الوقائع اللغوية المدروسة أشكالها وأمثلتها وتعميماتها وتجريداتها بغض النظر عن الاختلافات التي يمكن أن تسمها لدى تحققها نطقا عند المتكلمين أو كتابة عند المؤلفين.

ثم إن هذه الأوصاف النسقية والصورية مرتبطة ارتباطا وثيقا بالنظرية اللسانية على ما تتمتع به من جهاز مفاهيمي وأسس إبستمولوجية وإيديولوجيا معينة تساهم كلها في توجيه منهجها وتبني وجهات نظر محددة في دراستها للغة وكيفية اكتسابها وتوليدها للنصوص

---

[1] في النص الفرنسي الأصلي، يستخدم المصطلح ***Informatique*** عوضا عن علوم الحاسوب ***Sciences de l'ordinateur***. وعليه أوردت مصطلح "اللسانيات المعلومية" في مقابل ***linguistique informatique***، مع العلم أن اللسانيات الحاسوبية المقابل لـــ***linguistique computationnelle*** هو مقترض (دخيل) إنجليزي مستخدم كذلك في الأدبيات الفرنسية، وبمعنى مغاير أحيانا.

[2] ينظر: www.ling.uqam.ca Glossaire, Service ATO (UQAM) et EBSI (Université de Montréal) (تاريخ التصفح 2012/07/10).

والجمل، وكيفية فكِّها لمختلف الرموز الصوتية وغيرها من أجل استنباط المعاني والأفكار والمقاصد والأغراض المتوارية فيها.

غير أن اعتماد اللسانيات المعلومية على الوسائل الحاسوبية في معالجتها لتلك الأوصاف النسقية والصورية المتصلة بالنظرية اللسانية حول اشتغال اللغة قد يثير عدة مشاكل منها تطويع اللغة على حساب الآلة، مما يؤدي إلى اختزال الكثير من جوانبها أو الإجحاف في التعاطي معها، أو قد يؤدي إلى التسرع في الحكم على الظاهرة اللغوية المدروسة أو على النظرية اللسانية التي يرجع إليها وصف الوقائع اللغوية.

إن تلك المشاكل المثارة يمكن أن يقلل منها اهتمام اللسانيات المعلومية بالمقاربات النظرية التي تأخذ بعين الاعتبار أبنية الجمل على أنها مواضيع معقدة بالأساس لها تمظهرها الخاص الذي يخفي في ثناياه منطقا تهدف تلك المقاربات النظرية إلى الكشف عنه. وحتى تكون تلك العملية موضوعية وذات قيمة علمية، تقوم تلك المقاربات النظرية بترجمة خطاباتها الطبيعية إلى خطابات اصطناعية ومنطقية بعيدة عن الغموض واحتمال المعاني المتعددة وغيرها من الظواهر التي تسم اللغة الطبيعية حتى يمكن تشغيلها حاسوبيا.

ومتابعة لما تقدم من تعريف "اللسانيات المعلومية"، هناك فريق آخر من الباحثين يجعلها فرعا من اللسانيات الصورية[1] يهتم بتشغيل النماذج أو المنوالات الخاصة باللغة، بعبارة أخرى: تدرس اللسانيات المعلومية مختلف الإجراءات التي تفعِّل أو تنشِّط النماذج اللسانية وتحولها من حالة السكون إلى حالة الحركة والنشاط، ويتم ذلك وينفذ في أغلب الأحوال عبر برمجيات

[1] ليست اللسانيات الصورية فرعا لسانيا مستقلا بقدر ما هي اتجاه لساني يعمل على صورنة الظواهر اللغوية من أجل استخراج الخصائص الشكلية المشتركة بين جميع اللغات البشرية.

حاسوبية. وهو تخصص يختلف عن المعلوميات اللسانية التي تعنى بكتابة برمجيات حاسوبية تحاول لدى تنفيذها تقييس أو نمذجة بعض السلوكات اللغوية، كما يمكن أن تحيل على كل البحوث في المعلوميات التي تعمل على تطوير قدرة الحاسوب على تخزين المعلومة اللغوية أو بصفة عامة على تشغيل المعلومة اللغوية[1].

وعليه، فإن الكثير من الباحثين يفرق بين اللسانيات المعلومية وبين المعلوميات اللسانية، فليس لهما الارتباط نفسه باللسانيات؛ إذ أن المعلوميات اللسانية تأسست حول لسانيات محدودة تتمثل في الجانب الصرفي–التركيبي، فأنتجت محللات صرفية وتركيبية. بالمقابل، تكون المعلوميات في خدمة اللسانيات المعلومية لتلبية حاجات لغوية، لذلك يجدر تسميتها باللسانيات الموجهة بالحاسوب[2]. من هنا تظهر أوجه الاتفاق والاختلاف بين التخصصين؛ كلاهما يستخدم تكنولوجيا الحاسوب لمعالجة المعلومة اللغوية ولكن بأشكال وأهداف متباينة.

في اللسانيات المعلومية، ينطلق الباحث من أسس نظرية مبنية على النظر في طبيعة اللغة وكيفية اشتغالها، وعلى تصور أساليبها المنتجة للملفوظات أو المنشئة للمعاني والدلالات، ثم يصوغ ذلك النظر وهذا التصور في نموذج أو قالب يمثل تلك الخصائص والأساليب كلها، ولا ينتهي عمل الباحث حتى يتأكد من صلاحية ذلك النموذج بعرضه على التجريب والتمحيص والسبر مستخدما تكنولوجيا الحاسوب، وبالتالي فإن كتابة البرمجيات الحاسوبية هنا ليست هدفا في حد ذاته بقدر ما هي وسيلة لتجريب النماذج التي تبدو للباحث أنها تحاكي وتماثل سلوكاتنا

---

[1] ينظر: Polguère (A.), *Structuration et mise en jeu procédurale d'un modèle linguistique déclaratif dans un cadre de génération de texte*, thèse de doctorat, Université de Montréal, 1990, p. 1.

[2] ينظر: Rastier (François) et *al.* Sémaaaaalyse i i i ii iiieeeeeerrrrrrr rrrrrrr *op.cit*, p. 2.

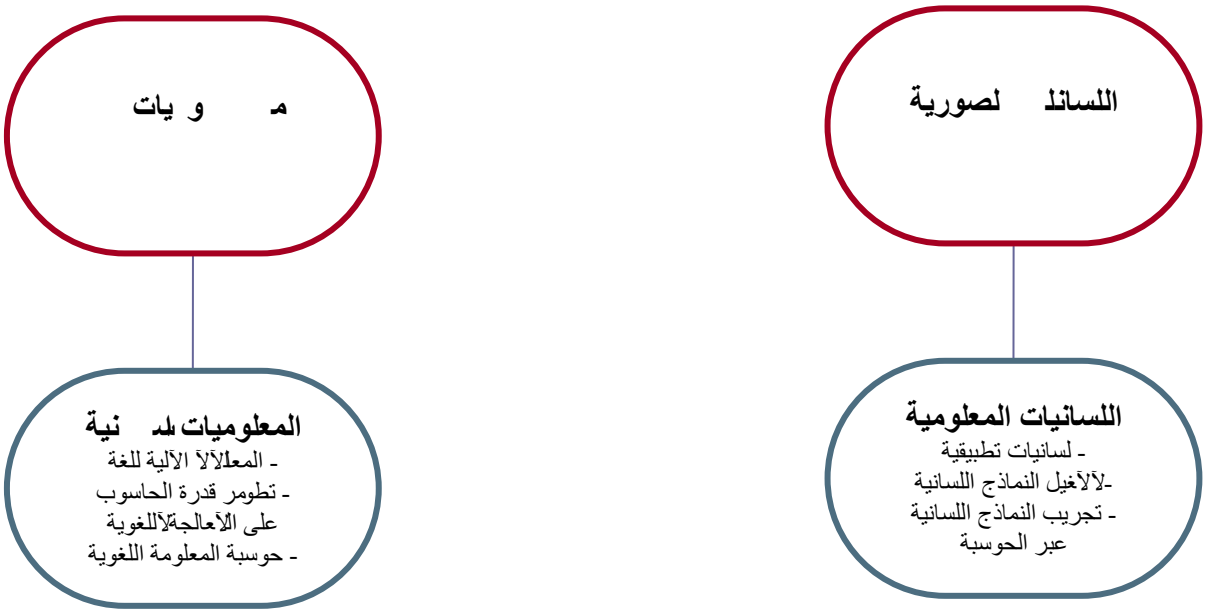
اللغوية. أما في المعلوميات اللسانية، فنجد كتابة البرمجيات الحاسوبية غاية في حد ذاتها، ولا يشترط أن يعتمد النموذج فيها على أساس لساني صريح.

اعتمادا على رأي **فوك Fuchs**، يظهر أن المعلوميات اللسانية هي فرع المعلوميات الذي يتناول معطيات لسانية، ومناهجُ المعالجة وتقنياتها هي نفسها المستخدمة في سائر ميادين المعلوميات، ولكن في هذه الحالة هي مطبقة على معطيات لسانية. كما يحيل هذا المصطلح أيضا إلى استخدام البرامج من أجل إقامة حسابات حول الكلمات وسلاسل الكلمات الموجودة في النص. بعبارة أخرى: المعلوميات اللسانية تعني مجموعة المعالجات الآلية للمعطيات اللسانية، مما يجعلها قريبة المعنى من المعالجة الآلية للغات.

أما اللسانيات المعلومية أو الحاسوبية، فهي فرع اللسانيات تستعمل أدوات طورتها المعلوميات من أجل تحقيق فرضيات نظرية لسانية حول اشتغال اللغة. لكن الملاحظ عند فرق البحث استعمالهم اللسانيات المعلومية للتعبير عن كل الأعمال في المعالجة الآلية للغات التي تعتمد بشكل أو بآخر على تحليل العناصر اللغوية، مما يؤدي إلى تطابق في استعمال المصطلحين "اللسانيات المعلومية" و"المعلوميات اللسانية"[1].

---

[1] ينظر: Fuchs (Catherine) et *al*., Linguistique et traitement automatique des langues, Hachette, Paris, 1993, pp. 22.



الشكل5: التقابل بين اللسانيات المعلومية والمعلوميات اللسانية

في الأدبيات الإنجليزية، هناك مصطلحان يستخدمان للدلالة على المجال الذي يجتهد فيه الباحثون في ابتكار البرامج الحاسوبية لمعالجة اللغة في مجال الذكاء الاصطناعي، هما ***Computational Linguistics***، و***Natural Language Processing***. يوظَّف المصطلح الأول عادة عند الباحثين المنتمين إلى **الاتجاه الهندسي** ***Engineering Approach***، أما المصطلح الثاني فيفضل استعماله الباحثون الذين ينحون منحى **معرفيا** ***cognitive-science Approach***.

كلا الفريقين متفقان في رغبتهما في جعل الآلات تؤدي مهام لغوية من خلال مختلف البرامج الحاسوبية، لكن افترقت لديهما الغاية؛ فأحدهما، وهو الاتجاه الهندسي، يعمل على ذلك لأنه يفضل المزيد من الآلات التي تحقق الفائدة للإنسانية نظرا لما تتمتع به من قدرة على تخزين المعلومات وسرعة معالجتها. وثانيهما، وهو الاتجاه المعرفي، لديه تلك الرغبة أيضا

لأنه يرى هذه العملية وسيلة للكشف عن كيفية قيام الإنسان بتلك المهام خاصة أنه يتعذر عليهم الجزم والبرهان على كيفية قيامه بتلك المهام في الواقع.

رغم هذا الافتراق، تبقى التقنيات المستخدمة في ذلك متقاربة، فمن أجل كثير من الخدمات التي يريد الذكاء الاصطناعي الهندسي أن يجد لها حلا تعتبر الأنظمة الإنسانية هي الوحيدة القادرة على تنفيذ تلك الخدمات، لذلك لا يمكن القيام بتلك الخدمات إلا من خلال محاولة محاكاة أو تقييس ما يعرفه الباحثون عن تلك الأنشطة الإنسانية. وعلى الجهة المقابلة، من أجل كثير من الخدمات التي يريد الذكاء الاصطناعي المعرفي أن يجد لها حلا، يقوم الباحثون في هذا الإطار باقتراح نماذج حاسوبية ثم ملاحظة ما يتلاءم منها مع الكفاية اللغوية عند الإنسان[1].

2.1. اتصال اللسانيات بالنظريات الحاسوبية

تتميز الأربعينات من القرن العشرين بانتشار البحوث الحديثة في مجال الإعلام والاتصال خاصة مع **شانون Shannon** و**ويفر Weaver** من خلال نظريتهما الرياضية للمعلومة حيث اعتبرتها معطى إحصائيا بالإمكان تكميمها وقياسها ونمذجة العملية التي يحدث فيها بث المعلومة واستقبالها، واستطاعت هذه النظرية أن تتكيف مع الحواسيب الحديثة النشأة، وأن تؤثر في العلوم المجاورة لها ومنها اللسانيات[2].

بالتزامن مع ذلك، انصب اشتغال الباحثين اللسانيين على نمذجة العمليات اللغوية حيث عمل أمثال **هاريس Harris** و**هوكيت Hockett** على تطوير بنية رياضية للغات الطبيعية،

---

[1] ينظر: Malmkjær (Kirsten), The linguistics encyclopedia, *op.cit*, pp. 28-29

[2] ينظر: Lazar (Judith), La science de la communication, Editions Dahlab, Alger, 2e éd., 1993, p. 36.

تتمثل في صياغة نموذج تحويلي يقوم بإنتاج الجمل انطلاقا من جمل نووية[1]. كما عمل تشومسكي أيضا بالموازاة مع مهندسي الإعلام والاتصال على تطوير إوالية لمجموعة متناهية من الحالات، وهي تشبه الآلة الحاسبة حيث يتم فيها الانتقال من حالة أولية إلى حالة نهائية مرورا بحالات وسطى، وفي كل مرة يجري استبدال رمز بآخر.

إن هذه الإوالية ذات الحالات المتناهية تقوم في كل معبر أو انتقال ببعث رموز معينة كمورفيم مثلا، وكل انتقال من حالة إلى أخرى يتوافق مع تعليمة ترسلها الإوالية كقاعدة نحوية مثلا. تنطلق الإوالية من حالة أولية وتمر على التوالي بسلسلة من الحالات وهي تبعث بمرفيم في كل انتقال حتى تصل إلى الحالة النهائية. يمكن القول: إن السلسلة المنتجة بهذه الطريقة تشكل الجملة، وكل إوالية من هذا النوع تحدد لغة أي مجموعة سلاسل المورفيمات التي يمكن إنتاجها بهذه الطريقة[2].

يمكن أن نمثل لذلك بالرسم البياني التالي الذي يتضمن نحوا يولد الجمل:

- الطفل يبدو مهذبا.
- الولد يبدو مهذبا.
- الطفل يبدو متخلقا.
- الولد يبدو متخلقا.
- الطفل يبدو محبوبا.

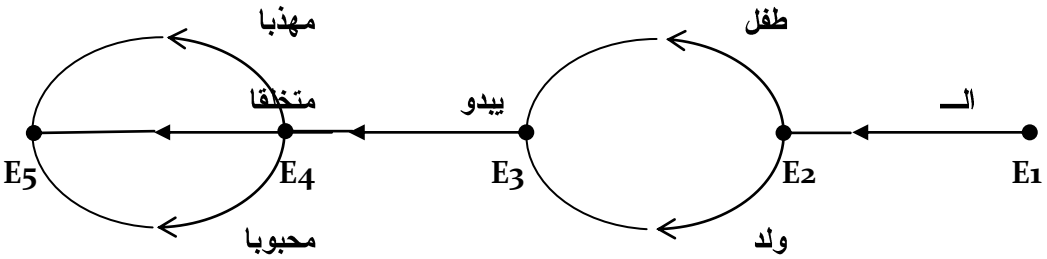
---

[1] ينظر: Mounin (Georges), « Le modèle en linguistique », Encyclopædia Universalis, Paris, 2002, tome XV, p. 304.

[2] ينظر: Saaa G Gaa aaa a e eeeeeeeeee eeeeeeeeeeeeage : 2. Processus de compréhension, Hermes, Paris, 1989, pp. 74-75.

- الولد يبدو محبوبا.

E1 إلى **E5** تمثل حالات الانتقال في توليد الجمل من الحالة الأولية إلى الحالة النهائية مرورا بالحالات الوسطى.



الشكل6: نموذج الحالات المتناهية

بالنتيجة، إن الاتصال بين اللسانيات والحاسوبيات يتجلى بشكل واضح مع النحو التوليدي، وقد أكد تشومسكي هذا الاتصال حين اعتبر نسق القواعد الذي يشكل بنية النموذج التوليدي التحويلي نسقا تضبطه النظرية الحاسوبية؛ يقول تشومسكي: « فاللغة –باختصار– يبدو أنها في جوهرها نظام حوسبي غني معقد البنية بدقة كاملة، وصارم في عملياته الأساسية»[1]، إن ذلك يعني أن القواعد النحوية عند المتكلمين تشبه إلى حد كبير القواعد الصورية التي يتبعها الحاسوب أثناء أدائه للحسابات.

يرجع ذلك التقارب بين النظريات اللسانية والنظريات الإعلامية إلى سعيها لتحقيق أهداف علمية وتكنولوجية، تتمثل أهدافها العلمية في الإسهام في معرفة بنية العقل البشري

[1] تشومسكي (نوم)، المعرفة اللغوية: طبيعتها وأصولها واستخدامها، تر. محمد فتيح، دار الفكر العربي، القاهرة، ط.1، 3993، ص. 707.

وطريقة عمله من زاوية لسانية، أما أهدافها التكنولوجية فتتضمن بناء حاسوب يحاكي قدرات الكائن البشري اللغوية، وتوظيفه في مجالات تطبيقية متعددة.

بالإضافة إلى ذلك، يعتبر الحاسوب وسيلة من وسائل تقييس النماذج المعرفية لأن المشتغلين بالبرمجيات في إطار الذكاء الاصطناعي ذلك المجال الذي يحاول فيه المتخصصون جعل الآلات تؤدي مهام تعتمد عادة على العقل[1]، يتقاسمون مع الباحثين اللسانيين وعلماء النفس المعرفيين وعلماء الأعصاب البحث في آليات اشتغال الدماغ عند الإنسان، وكيفية حله للمشاكل ومعالجته للمعارف التي تأتيه من المحيط، وكيفية قيامه بالتخطيط والعمل واتخاذه للقرارات.

وعليه، تمت استعارة هذا المبدأ في علم النفس المعرفي في تصورها العام للمعمارية الذهنية للمعلومات وشبهت بآليات المعالجة في الحاسوب الذي يستقبل المعلومات ويقوم بتشفيرها ويعالجها في سلسلة من المراحل والمستويات، إما بطريقة متسلسلة أو متوازية، ثم يخزنها في الذاكرة المركزية ويستعملها في تنفيذ المهام. كما يتكون الحاسوب من مكونين هما الجهاز المادي **hardware** والجهاز البرمجي **software** الذي يمكنه معالجة المعلومات بشكل مستقل عن الجهاز المادي، حيث يمكن الجهاز المادي أن ينجز عمليات حسابية انطلاقا من توجيهات الجهاز البرمجي.

في مقابل استفادته من الحاسوبيات ونظرية الإعلام، كان للنحو التوليدي دور كبير في تطوير الدراسات الحاسوبية من خلال البحث في خصائص اللغات الصورية ومدى كفايتها في بناء الأوصاف التركيبية والدلالية للغات الطبيعية، وتأسيس لغات البرمجة الحاسوبية التي تسهل

---

[1] ينظر: Malmkjær (Kirsten), The linguistics encyclopedia, *op.cit*, pp. 28-29

على مستخدمي الحاسوب التواصل معه[1]، وكما استلهم الحاسوبيون حصيلة أبحاث اللسانيين لتطوير خوارزمات تقبل الإدماج في برمجيات حاسوبية صممت لأجل إنجاز معالجة آلية للغات الطبيعية، وتوفير المحللات الصرفية والتركيبية الملائمة لوصف اللغة الطبيعية[2].

3.1. منهج اللسانيات الحاسوبية

يبدو لنا مما تقدم أن مصطلح "اللسانيات الحاسوبية" هو مصطلح عام يشمل كل تلك المجالات المتقدمة والتي تتداخل فيها اللسانيات بعلوم الحاسوب حيث يتناول فيها الباحثون موضوعات لغوية باستخدام الحاسوب، مهما اختلفت المناهج والمقدمات والتوجهات العلمية بين هندسية يطغى عليها التوجه الآلي التقني وبين معرفية يشيع فيها إنشاء النماذج الحاسوبية، سواء في جانبها النظري أو من خلال تطبيقاته المتعددة.

وبالتالي، يختلف الباحثون في تحديد مناهج اللسانيات الحاسوبية تحديدا واضحا، ولعل ذلك راجع إلى تجاربهم ومشاربهم العلمية؛ فإن كان الجميع متفقون على أن هذا العلم يعالج المواد اللغوية في الآلات الإلكترونية، فإن بعضهم يجعله جزءا من الذكاء الاصطناعي، وبالتالي يفرض على اللسانيات الحاسوبية مناهج الذكاء الاصطناعي. هذا الأخير الذي يتميز بجوانب نظرية وأخرى تطبيقية. تتضمن الجوانب النظرية معرفة الإطار النظري العميق الذي يعمل في الدماغ البشري لحل المشكلات الخاصة كالترجمة من لغة لأخرى، وتتضمن الجوانب التطبيقية

---

[1] ينظر: Yousfi (Abdellah), Traitement automatique de la langue (texte et parole), Bouregreg, Rabat, 1re éd., 2006, p. 13.

[2] ينظر: علوي (حافظ إسماعيلي) والملاخ (امحمد)، قضايا إبستمولوجية في اللسانيات، منشورات الاختلاف، الجزائر، ط.1، 2009، ص. 104.

التعامل مع الرياضيات الخوارزمية والتي هي مجموعة من القواعد ترتب بشكل معين لتعطي نتائج مماثلة للنتائج التي نجدها لدى الإنسان.

غير أن هناك باحثين يربطون اللسانيات الحاسوبية بحقل الإحصاء اللغوي للمواد اللغوية، وهنا يعتمد الباحث على المناهج الإحصائية لحل المشكلات، وفريقا ثالثا من الباحثين يرى بأن اللسانيات الحاسوبية هي تصميم وتطبيق لتقنيات العمليات الرياضية الخوارزمية بهدف تحليل اللغات البشرية وتركيبها، وهي في ذلك متصلة باللسانيات العامة تستمد منها المفاهيم الأساسية حول اللغة وكيفية اشتغالها، وبالذكاء الاصطناعي من حيث إنها بحاجة إلى تمثيل[1] مختلف المعارف اللسانية النحوية منها والدلالية في الحاسوب على غرار ما تمثل به في الدماغ البشري[2].

إن المعطيات التي تجري عليها عملية الحساب في الرياضيات الحاسوبية هي كميات معلومة تكون على شكل كائنات رياضية، وكذلك الحال في الفيزياء الحاسوبية فإن معطياتها الحسابية هي كائنات فيزيائية. وهذا ما يقال أيضا عن اللسانيات الحاسوبية التي تتخذ من المعطيات الحسابية موضوعا لها[3].

الملاحظ إذن حسب ما مر بأن طبيعة المعطيات المتناولة هي تابعة للنظرية ومتوقفة عليها، فقد ينطلق الباحثون من نظريات رياضية أو فيزيائية أو لسانية وتكون معطياتها تبعا

---

[1] التمثيل هو المعطى الأولي للتفكير، وهو المادة الخام لبناء المعرفة، والإدراك الحسي لا يفهم إلا بواسطة التمثيلات. ينظر: أدهم (سامي)، إبستمولوجيا المعنى والوجود ونقد التطورية، مركز الإنماء القومي، ص. 15.

[2] ينظر: الوعر (مازن)، قضايا أساسية في علم اللسانيات الحديثة، دار طلاس، دمشق، ط. 1، 1988، ص. 777-119.

[3] الحساب هو فن تنفيذ العمليات بواسطة نظام علامات أو نسق دوال. هناك فرق بين الحساب والأرثميتيقا التي هي الأعداد ذات القيم المحدودة والمعبر عنها بالأرقام. ينظر: Cuvillier (Armand), Nouveau vocabulaire philosophique, Armand Colin, Paris, 1956, pp 22, 29 et 58.

لذلك المنطلق. ومنه، فإن تخصيص تلك العلوم بالحاسوبية لا يجعل منها علوما جديدة بقدر ما يسلط الضوء على المنهجية المطبقة في تلك العلوم؛ إذ لا يقوم الباحث فيها بمحاولة مفهمة[1] المعارف المتصلة بالمعطيات أي لا يعمل على استخراج المفاهيم الواصفة للظواهر الملاحظة، بل يحاول مفهمة المعارف المتعلقة بحساب تلك المعطيات أي يعمل على صياغة المفاهيم التي تحيل إلى انتقال تلك المعطيات من حال إلى حال، أو بتعبير آخر: يعمل على صياغة مفاهيم وتصورات حول اشتغال تلك المعطيات. كما أن صفة "الحاسوبية" لا ترادف "العملية" ***pratique*** التي تشمل كل ما يتعلق بالفعل والعمل في مقابل "النظرية" والتأمل الفلسفي.

وبالتالي، فإن اللسانيات الحاسوبية ليست علما جديدا مستقلا عن اللسانيات النظرية، بقدر ما هي تخصص فيها يتبنى منهجا خاصا في دراسة الوقائع اللغوية يتمثل في دراسة النواحي الحاسوبية لدى إنتاج اللغة وتحليلها، وذلك من أجل إنشاء البرامج الحاسوبية غالبا.

لذلك ينبغي أن يكون للاستخدام الحاسوبي رافد نظري من العلم نفسه الذي نريد حوسبته، وبالتالي يكون الاعتماد على الأسس النظرية التي يتيحها المجال المعرفي الذي ينكب الباحثون على دراسة اشتغال معطياته وإقامة التصورات حول العمليات التي تجري عليها. وعليه، ينبغي الرجوع إلى الأسس النظرية التي وضعتها اللسانيات العامة والاستفادة منها في إثراء البحث اللساني الحاسوبي والذي هو ليس بمعزل عنها[2].

---

[1] المفهمة هي اختزال المعطيات التجريبية في مفاهيم، من المحسوس إلى المعقول: "العلم هو مفهمة للطبيعة". ينظر: Ibid., p 39

[2] ينظر: Ratté (Sylvie), *Interprétations des structures syntaxiques : une analyse computationnelle de la structures des évènements*, *op.cit*, p. 49.

إن تبني مقاربة حاسوبية من أجل دراسة موضوع معين يعني دراسة هذا الموضوع ضمن ثلاث مستويات من التجريد: مستوى المكنزم، والمستوى الخوارزمي، والمستوى الحاسوبي المحض، هذه المستويات تمثل الخواص الآتية[1]:

(1)	***مستوى المكنزم***: وصف المهام التي تقوم بها العناصر المادية للدماغ (أو الحاسوب).
(2)	***المستوى الخوارزمي***: وصف الخوارزمية التي تتحكم في نشاط الجهاز. تتيح هذه المقاربة عدة تشغيلات ممكنة وغير محدودة تقريبا بما أنها مرتبطة بالجهاز المتوفر.
(3)	***المستوى الحاسوبي***: أعلى مستويات التجريد، يتعلق بتحليل المشكل في إطار معالجة المعلومة (أي النقل الرمزي للمعلومة) بمعنى يتم تحديد ما يمكن حسابه ولماذا، وكذا توفير نماذج رياضية لهذه المشاكل.

الجدول5: مستويات التجريد في الدراسات الحاسوبية

يعتمد البحث في اللسانيات الحاسوبية على النمذجة والتي هي مجموعة من الفرضيات حول مستويات التمثيل يفسر بها الباحث مظهرا حاسوبيا للغة ويطور فيها نماذج نظرية لتفاعل تلك المستويات، وتظهر جليا في المستوى الثالث من مستويات التجريد في الدراسة الحاسوبية للغة. إذن النظرية اللسانية الحاسوبية تتبع النمذجة، وهي نظرية للحسابات اللسانية، تضع

[1] ينظر: *Ibid.*, p. 54

فرضيات وتقترح مبادئ وقيودا وقواعد لتلك الحسابات ولمواضيعها. بصفة عامة، إن أطروحة اللسانيات الحاسوبية تفترض وجود معجم يمكن أن توصف فيه بنية السيرورات الذهنية الخاصة باللغة أي اللغة الضرورية للتعبير عن الخوارزميات[1].

إن أهم ميزات اللسانيات الحاسوبية هي النمذجة الحاسوبية حيث تضم هدف النمذجة وهو وصف السيرورات التي يمكن حسابها والتي يتم فيها تركيب اللغة وتحليلها، بالإضافة إلى تحليل المشاكل في مستواها الحاسوبي، والمقصود بالنمذجة صياغة نماذج أي أنظمة عملياتية تحاكي بنيتها العلائقية سيرورة معينة كسيرورة اللغة.

لقد شهد نشاط النمذجة في اللسانيات تطورا معتبرا بدءا بسنوات 1970. تهتم النماذج اللسانية بالظواهر المشاهدة، مثل كل العلوم التي تعالج معطيات تجريبية. يتعلق الأمر بوضع تصور لجهاز يعطي تشغيله نتائج مشابهة لتلك المعطيات اللغوية المشاهدة. النمذجة إذن وسيلة لتشغيل النظرية؛ تسبر مدى مناسبتها للوقائع التي تعمل على تفسيرها. من هذا المنظار، كلما كان النموذج بسيطا، كانت قدرته التفسيرية أكبر، بما أنه يمكن تعيين دور كل عنصر من النظرية أثناء اشتغاله، لكن هذه البساطة المثالية تتماشى مع ضرورة الموافقة القصوى للمعطيات في أكبر دقة ممكنة، مما يدعو إلى تعقيد النموذج. إن صياغة النماذج تتم بمشاركة الرياضيات بصفة مباشرة أو عبر المعلوميات، فكل نمذجة حاسوبية تتفرع دائما عن نموذج رياضي تفرعا ظاهرا أو مقدرا[2].

---

[1] ينظر: Pylyshyn (Zenon W.), « Computation and cognition: issues in the foundations of cognitive science », The Behavioral and Brain Sciences, USA, 1980, N°3, p. 116.

[2] ينظر: Mounin (G.) « Le modèle en linguistique », tome XV, Encyclopædia Universalis, Paris, 2002, p. 303.

من خلال ما تقدم من مناقشة لمصطلحات "اللسانيات الحاسوبية" وحقولها ومناهجها، تبين أن حقل المعالجة الآلية للغات الطبيعية هو من أثراها وأكثرها عناية عند الباحثين، وهو يشكل المجال التطبيقي التي تتمازج فيه النظريات اللسانية والحاسوبية ووتتلاحم من أجل أن تتمخض عنها دراسات تصل المعطيات اللغوية بالجوانب الحاسوبية، ويتولد عنها برامج حاسوبية متطورة قدر الإمكان حتى تعالج مختلف المهام والخدمات اللغوية، لذلك أرى من اللازم أن أخصص لها الفقرات التالية من هذا الفصل.

2. المعالجة الآلية للغات الطبيعية

إن المعالجة الآلية للغات هي حقل بحث متعدد الاختصاصات، يساهم فيه لسانيون وحاسوبيون ومناطقة ونفسانيون ووثائقيون ومعجميون ومترجمون، ويندرج في ميدان الذكاء الاصطناعي. لقد أنشأ المهندسون إواليات يمكن أن تلعب الشطرنج، أو تسوق السيارات أفضل من الإنسان أحيانا، لكن تألية الأنشطة اللغوية الإنسانية، أي جعلها تتم بطريقة آلية، ما زالت تعاني من عوائق، ولا يمكن تحقيقها 100%. ومع ذلك، هناك بعض المهام المحددة التي تستخدم اللغة في المعالجة الآلية يمكن أن تعطي برمجيات جيدة، لكن في أغلب الأحيان تحتاج ضرورة للتدخل البشري قبل عملية المعالجة (من خلال تبسيط الأوصاف مثلا) وبعدها (بالتصحيح مثلا) من أجل الحصول على نوعية مماثلة لما ينتجه الإنسان.

بالإضافة إلى ذلك، تعد المعالجة الآلية للغات واقعا اجتماعيا واقتصاديا، بما تتيحه من مؤسسات ومنتوجات متخصصة: مصححات الإملاء، برمجيات الترجمة، برمجيات الإملاء الصوتي... ومع الحوسبة المتصاعدة للأنشطة، لا تتوقف الميادين التطبيقية للمعالجة الآلية

للغات عن التوسع مع تفاوت بينها في الأهمية: فمستواها ضعيف مثلا بالنسبة للسيارات التي تتكلم، ومتوسط بالنسبة لبرمجيات معالجة النصوص التي تتضمن مصححات إملائية ونحوية، ومعتبر بالنسبة للتطبيقات الذكية كالترجمة الآلية[1].

تشكل المعالجة الآلية للغات أيضا تحديا سياسيا وثقافيا بالنسبة للدول؛ ففي الوقت الذي تنفتح هذه الدول على الثقافات المختلفة واللغات الأجنبية في زمن العولمة، يأتي اجتياح البرمجيات المصممة بتلك اللغات خاصة الإنجليزية والفرنسية ليؤدي إلى اختلال في توازن اللغات المحلية، والعربية ليست بمنأى عن هذه الخطورة، فحتى لو كانت معظم معالجات النص متكيفة مع اللغة العربية غير أنه ما يزال الكثير من البرمجيات المتخصصة، وبرمجيات الألعاب والتواصل الاجتماعي في الإنترنت باللغة الإنجليزية، وكذا البرمجيات التعليمية الموجهة بالحاسوب التي تعرض المواد التعليمية باللغة الفرنسية، حيث تفرض هذه البرمجيات على مستخدمها نظاما ثقافيا أجنبيا عنه[2].

إن المعالجة الآلية للغات هي بالمعنى الحصري ميدان جديد ومعاصر مرتبط بظهور الحاسوب، وهو يدل في عمومه على محاولة لتقييس الأنشطة اللغوية عند الإنسان (إنتاج التقارير، تحليل النصوص، القراءة، الكتابة، الترجمة، تصحيح الأخطاء الإملائية والنحوية، تلخيص النصوص...) باستخدام وسائل ميكانيكية. وقد ورث الحاسوب ذلك عن الطباعة والآلة الكاتبة اللتين تمت فيهما تألية نسخ النصوص وكتابتها، وقد ساعدت على ذلك أيضا تلك النتائج المذهلة في تشفير الكلام ونقله باستخدام الهاتف.

---

[1] ينظر: Abeillé (Anne), « Traitement automatique des langues » *in* Encyclopædia Universalis, 2002.

[2] ينظر: Fuchs (Catherine) et *al*., Linguistique et traitement automatique des langues, op.cit, p. 14.

وبداية من سنوات 1950 وفي سياق الحرب الباردة، قام الباحثون بتطوير الحاسوب بالتوازي مع استخدامها في مجال معالجة اللغات، فازدهرت الترجمة الآلية خاصة بين اللغتين الإنجليزية والروسية، لكن هذه العملية التي تعتمد على الترجمة الحرفية للنصوص تفتقر كثيرا إلى المعارف اللسانية ونتائجها كانت مخيبة للباحثين[1]. وبين سنوات 1960-1980، استقلت المعالجة الآلية للغات عن نظريات الإعلام لتدخل في حيز الذكاء الاصطناعي. وعليه، تم تطوير أنظمة خبيرة للتعرف على الأشكال الصوتية وإنتاجها موازاة مع تلك المخصصة للأشكال المرئية، وبالنسبة للكتابة تم تطوير برمجيات التحليل بالتزامن مع تطور النحو التوليدي والصورنة اللسانية[2].

انطلاقا من سنوات 1980، شُرع في طور التصنيع بصناعة أنظمة أقل طموحا لكنها أكثر فعالية، فكان التركيز أكثر على سرعة عمليات الحساب وعلى تنظيم الموارد اللسانية كالمعاجم والأنحاء الإلكترونية، وصار التركيز أقل على صورنة عمليات الحساب التي تشبه أكثر فأكثر الحجر السوداء. فنتجت عن ذلك تطبيقات كثيرة وغير منتهية تقريبا في ميادين متنوعة: المكتبيات، الوثائقيات، الترجمة، التعليم، الصحافة، النشر، الاستعلامات، المعلوميات الطبية، مساعدة ذوي الاحتياجات الخاصة...

لقد أدى هذا التطور المتزايد في التطبيقات المتنوعة للمعالجة الآلية للغة إلى نشأة ما يسمى **بالصناعات اللغوية *industries de la langue*** التي تضم مجموعة من الأعمال التي

---

[1] ينظر: Yousfi (Abdellah),Traitement automatique de la langue (texte et parole), op.cit, p.14.

[2] ينظر: Abeillé (Anne), « Traitement automatique des langues » *in* Encyclopædia Universalis, Paris, 2002.

تشارك في إنشاء برمجيات تجارية تعالج معطيات لغوية، وذلك بطريقة إجرائية لكن ضمن مجالات محدودة. هذه البرمجيات التي تعتمد على قواعد معارف لسانية تتناول اللغة في بعدها الشكلي، ولا تشتغل إلا ضمن إطار دلالي ضيق كأنظمة بيانات ومعاجم متخصصة وبرامج تعليمية[1].

1.2. المفاهيم الأساسية

الهدف من المعالجات الآلية للغات هو ابتكار برامج حاسوبية بإمكانها أن تعالج آليا البيانات اللسانية، أي معطيات معبرة بواسطة لغة "طبيعية". لكن: ماذا نقصد بتلك المفاهيم الثلاثة؟

يتناول هذا المجال دراسة اللغة من خلال معطياتها اللسانية التجريبية، فهذه المعطيات هي ذات أشكال متنوعة: إذ يمكن أن تكون نصا مكتوبا أي سلسلة من الجمل ذات محتوى اتصالي متسق، أو حوارا شفويا أو مكتوبا، أو وحدات لسانية أصغر حجما من النص كالجملة أو الملفوظ أو الكلمة... لكن الباحثين يشيرون إلى كل تلك الأشكال عادة بمصطلح "النص" ويجعلونه موضوع دراستهم.

أثناء عملية المعالجة، يتناول الحاسوب موضوع الدخول بتغييره إلى موضوع آخر هو موضوع الخروج. وبحسب طبيعة التطبيق، تتم المعالجة إما بتحويل نص إلى نص آخر كإعطاء نص جديد مصحح من النص الأول، أو استخراج المعلومات منه، أو تقليصه، أو

[1] ينظر: Fuchs (Catherine) et *al*., Linguistique et traitement automatique des langues*, op.cit*, p. 15.

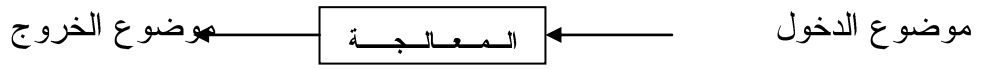
ترجمته...، وإما أن تتم المعالجة بصياغة نص غير موجود سلفا بالاعتماد على معلومات أولية أو أشكال صورية.

في حالة التحويل، يكون موضوع الدخول نصا **T** مكتوبا أو شفويا، أما موضوع الخروج فيمكن أن يكون تمثيلا لهذا النص **R(T)** يظهر مختلف المعلومات التي يتضمنها **T** ليوظف فيما بعد في تطبيق معين كالبحث عن المعلومات في قاعدة بيانات، كما يمكن أن يكون موضوع الخروج نصا جديدا **T'** يحمل نفس مضمون النص الأول بصياغة مغايرة، أو يلخصه أو يترجمه إلى لغة أخرى بعد أن يتم استخراج تمثيله في مرحلة وسطى.

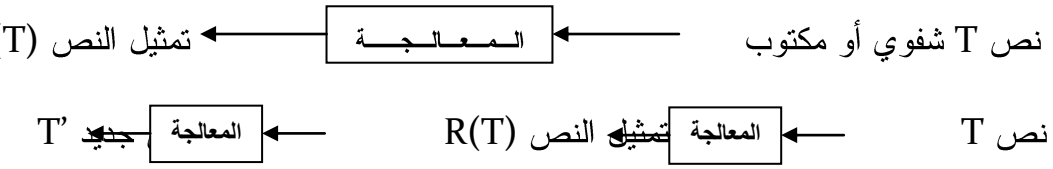
وفي الحالة الثانية التي تتضمن صياغة النص، يكون موضوع الدخول تمثيلا لنص يراد صياغته أي تمثيلا لمحتوى نريد التعبير عنه بواسطة نص **T** يشكل موضوع الخروج. هذا التمثيل الذي يحمل معلومات النص الذي نريد إنشاءه قد يأتي من مصدر خارجي عن برنامج المعالجة مثلا عبر برنامج آخر مهمته البحث عن المعلومات داخل قاعدة بيانات، أو برنامج يحتوي على بنك للصور...

يمكن أن نلخص أنواع المعالجة في المخطط التالي:

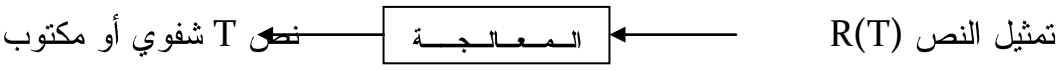
الحالة العامة:



الحالة 1: التحويل



الحالة 2: صياغة النص



وحتى تتم معالجة موضوع ما يجب معرفة أسس تأليفه الداخلية التي يبنى عليها، أي ينبغي على الباحث أن يتمكن من وصف بناء الموضوع (النصوص المشاهدة) بطريقة إجرائية؛ بعبارة أخرى: يوصف النص بأنه مجموعة من التوافقات بين الأشكال: الكلمات والجمل والفقرات من جهة، وبين المعاني: المدلولات والقضايا من جهة ثانية، وهذه التوافقات توجهها قواعد مفسرة هي قواعد اللغة، وبالتالي أثناء معالجة اللغة آليا يتعين على الباحث وصف اللغة التي ينتظم فيها النص.

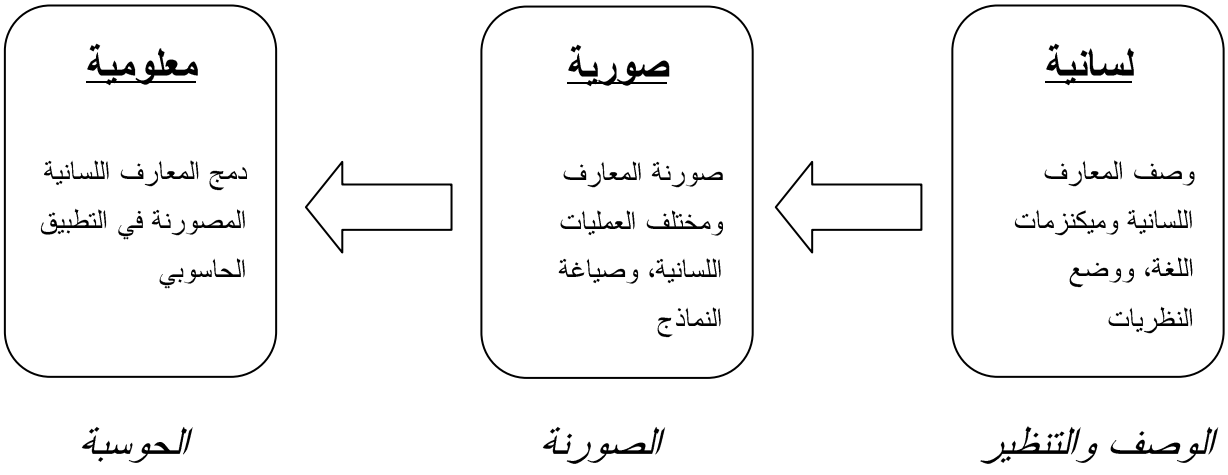
إن هذه المعالجة توسَم بأنها آلية لأنها تجري بوسائل ميكانيكية هي الحواسيب الآلية عادة، وهي أجهزة تقوم بالحسابات، وعليه فإن الباحث إذا أراد أن يستعين بتلك الأجهزة لإتمام عملية المعالجة آليا يحتاج أن يرجع المواضيع المعالجة إلى حسابات. وبالتالي، يمكن القول: إن المعالجة الآلية هي متتالية من الأفعال (الحسابات) التي تجريها الآلة (الحاسوب) وفق ترتيب زمني تعاقبي (البرنامج)[1].

بالإضافة إلى ذلك، يشترط لإنجاز عملية المعالجة الآلية توافر أدوات وتقنيات غير متجانسة: لسانية وصورية ومعلومية أو حاسوبية، نظرا إلى أن وصف ميكانزمات اللغة وملاحظتها يتطلبان التحكم في معارف وتقنيات بحث لسانية دقيقة، ولكن وصفها وملاحظتها في مجال المعالجة الآلية للغة يتطلب بالإضافة إلى ذلك الاستعانة بإجراءات صورية معقدة حتى

[1] ينظر: Fuchs (Catherine) et *al*., Linguistique et traitement automatique des langues, op.cit, pp. 7-9.

يتسنى التعبير عن المعارف اللسانية بواسطة أمثلات (صورانيات) تسهل إدماج تلك المعارف في الحاسوب. كما يحتاج الحاسوب إلى أن يزود بتقنيات واستراتيجيات معلومية تمكنه من أن يضع تلك المعلومات المصورنة تحت تصرف التطبيق المستهدف سواء أكان هذا التطبيق ترجمة آلية أم غيرها من التطبيقات الحاسوبية[1].

يمكن أن نجمل خطوات المعالجة وأدواتها في الشكل الآتي:



الشكل7: خطوات المعالجة وأدواتها

2.2. صعوبات المعالجة

إن الصعوبات التي تعترض البحث في ميدان المعالجة الآلية للغات قد تنبع من طبيعة اللغات وخصائصها، وقد تأتي من ظروف الاتصال. من بين الخصائص الداخلية للغات والتي تشكل صعوبات للمعالجة عددها الكبير وتنوعها، واتسامها بالغموض واحتمال عدة معان، وتغير الملفوظات في الواقع، وأخيرا الإبداعية اللغوية.

---

[1] ينظر: Fuchs (Catherine) et *al*., Linguistique et traitement automatique des langues, *op.cit*, pp. 11-12.

فيما يخص عدد اللغات، يحصي الباحثون أكثر من 5000 لغة إنسانية مستعملة حاليا في العالم، لكن إذا اقتصرنا على اللغات التي تملك أنظمة للكتابة والتي تشكل التأليفية فيها رهانا اجتماعيا واقتصاديا، لا يبقى إلا بضع مئات من اللغات التي يمكن أن تحظى بالمعالجة الآلية، ومع ذلك فليست كلها معروفة بتراث لساني ومعاجم وأنحاء يعتد بها، هذا ما يشكل عائقا أمام حوسبتها.

ثم هناك لغات تبدي مشاكل خاصة، مثل تفكك الرموز بالنسبة إلى اللغات غير الألفبائية كالصينية مثلا حيث تحيل هذه الرموز على أفكار وليس على حروف ألفبائية[1]، وكتابة الحروف في اللغات السامية خالية من الحركات... بالإضافة إلى ذلك، إمكانية تحقيق تطبيق حاسوبي في لغة معينة لا يستلزم بالضرورة إمكانية نقل هذا التطبيق ببساطة إلى لغة أخرى، وإن كان بعض البرمجيات يتكيف مع لغات متقاربة لكن ليس كلها. إن اللغة التي تحظى بالمعالجة الآلية ستستمر ويطول أجلها أكثر من غيرها، وحضورها في الإنترنت يدل على قيمتها ووزنها في العالم المعاصر[2].

أما بخصوص ظاهرة الغموض التي تطغى على اللغات الطبيعية، فإن كل لغة تملك عددا هاما من الغموض الكامن فيها، والتي تثير عدة فرضيات لفهم اللغة يمكن تخطيها من قبل الناس عند تحليلهم للملفوظات بسبب اعتمادهم على توجيهات غير لسانية كتقاسيم الوجه أو دلالة السياق، لكن الآلة تقف عاجزة عن ذلك. هذه الظاهرة التي قد تغيب عن وعي المتكلمين باللغة

---

[1] ينظر: Dubois, (Jean) et *al*., Dictionnaire de linguistique et des sciences du langage, *op.cit*, p. 238.

[2] ينظر: Théodule (M.-L.) et Gruszow (S.), «Quelles langues vont disparaître ?», La Recherche, n° 429, Avril 2009, p. 33.

لكنها حاضرة معروفة عند اللسانيين، ترجع أسباب ظهورها إلى الرغبة أحيانا في الاقتصاد في الجهد المبذول من قبل المتكلمين، لذا يلجأ إلى استخدام أشكال متماثلة للتعبير عن مدلولات مختلفة (الترادف)[1].

هذه الظاهرة تُلمح في كل المستويات: ففي المستوى الصوتي، قد ينطق الصوت اللغوي ويكتب بأحرف مختلفة مثل: الألف في عصا وأقصى، والهاء في مكتبة وماله... فيصعب التعرف الآلي على المنطوق وتسجيله. كما قد تنطق الكلمة ولها معان مختلفة (عين) مما يصعب التحليل الآلي والتأليف النطقي. كما أن الجملة نفسها قد تحيل إلى معان متعددة مما يصعب عملية الترجمة الآلية:

- زرت صديقي المريض أمس.

- *J'ai visité mon ami qui était malade hier.*
- *J'ai visité hier mon ami qui était malade.*

هذا فقط على مستوى المعاني المجردة المعجمية والصرفية والتركيبية، وإذا أضفنا إليها المقاصد والأغراض الخطابية والتداولية لكانت العملية أكثر صعوبة. يضاف إلى ذلك استخدام المصطلحات نفسها في مجالات علمية مختلفة قد يصعب على برمجية للبحث الوثائقي أن تستخرج النصوص المطلوبة، وبالتالي فالغموض لا يمكن تلافيه في اللغة.

يعتبر تغير الملفوظات وجها آخر للصعوبة التي تلاقيها المعالجة الآلية للغات. في التعرف الآلي على المنطوق مثلا، يقوم الباحثون بالبحث عن نطق متقارب للفونيم أو المقطع، والذي يتنوع بحسب الفونيمات المجاورة والمتكلمين رجالا ونساء، كبارا وصغارا والحالة

[1] ينظر: Abeillé (Anne), « Traitement automatique des langues » *in* Encyclopædia Universalis, Paris, 2002.

النفسية والجسدية لنفس المتكلم. كما أنه مصدر للصعوبة في التوليد والترجمة الآليين، حينما يحاول البرنامج اختيار الكلمة المناسبة أو البناء الأكثر ملاءمة للمعنى المعبر عنه، هذا الاختيار الذي يرتبط بمعايير صعبة الصورنة.

وأخيرا، اللغات تنمو، تظهر دوما مفردات وعبارات جديدة، وتختفي مفردات وعبارات أخرى، وتتحدد مضامينها بحسب الاتفاقات الاجتماعية. لذلك، يتعذر وضع قوائم قبلية للألفاظ ومدلولاتها، وللتراكيب ومعانيها بسبب هذه الإبداعية اللغوية. مهما يكن حجم المعجم الإلكتروني المنشأ، فهو يفتقر دوما إلى مفردات وعبارات غير واردة فيه، وإن كانت اللغة العربية من أكثر اللغات الحية ثباتا، وأكثرها حفاظا على معجمها وأنظمة تراكيبها. حتى لو زودت الآلة بكل المعارف اللسانية الضرورية، ينقصها دائما معارف أخرى غير لسانية نجدها في التواصل البشري مثل: الحركات المستخدمة أثناء الحوار، وكالمرجعيات الثقافية للمتخاطبين.

3.2. مقاربات المعالجة

لتجاوز تلك الصعوبات، اقتُرحت عدة استراتيجيات تتمثل خاصة في نمذجة المعارف اللسانية وغير اللسانية الضرورية، والنمذجة الرياضية التي تقوم على الانتظام المشاهد في النصوص، والنمذجة المعرفية التي تقوم على تقييس السلوكات البشرية.

- ***النمذجة اللسانية:***

في هذه الاستراتيجية، يجتهد الباحثون في توضيح المعارف اللسانية التي يمتلكها الفرد ويقومون بنقلها إلى الآلة. تتناول هذه العملية توزيع تلك المعارف على مستويات تتوافق مع الميادين اللسانية الفرعية (الصوتيات، الفنولوجيا، علم الصرف، علم التركيب، علم الدلالة،

التداوليات). لكن ينبغي اتباع نظرية معينة في مقاربة تلك المعارف اللسانية، يعتبر النحو التحويلي إحدى تلك النظريات الرائدة في مجال المعالجة الآلية للغة، حيث يعتبر الأنحاء عبارة عن خوارزمات تميز الصيغ السليمة البناء من غيرها، رغم ذلك فضل بعض الدارسين نماذج أخرى في وصف اللغة تعطي أهمية أكبر للمعجم والدلالة كالأنحاء التوحيدية (أنحاء التضام) أو أنحاء التعلق.

على العموم، تحقيق نظام للمعالجة الآلية للغات هو ثمرة بحوث اللسانيات الوصفية التقليدية وبحوث اللسانيات النظرية، حيث إن تراكم المعارف التجريبية حول اللغات يرجع للسانيات، أما الصورنة الضرورية في تشفير المعارف اللسانية من أجل برمجتها في الآلة يرجع إلى اللسانيات الرياضية. وفي تحديد اللغات الواصفة الخاصة بالبرمجة الحاسوبية هناك حركة ذهاب وإياب بين اللسانيات النظرية والمعالجة الآلية للغات.

وبالنسبة لنمذجة المعارف التركيبية والدلالية، يميز الباحثون بين مرحلتين: في المنطلق، المعالجة الآلية للغات استعانت في وصفها اللساني بلغات رياضية أو منطقية كانت حاضرة قبل ظهور هذا الحقل العلمي الجديد، مثل النحو الصوري لنمذجة علم التركيب، ومنطق المحمولات الذي يسمح بتناول الوظائف مستقلة عن قيمها في علم الدلالة. لكن بدءا بسنوات 1980، تم تطوير أدوات صورية (صورانيات أو أمثلات) أكثر تكيفا مع تمثيل مختلف المعارف اللسانية مثل الأنحاء التوحيدية.

إن الأنحاء التوحيدية ظهرت نتيجة لما اعترض النحو التوليدي التحويلي من مشاكل صعبت من إمكانية تشغيله الآلي[1]؛ فقد كانت البحوث اللسانية تصف التوليدات الممكنة من البنية العميقة عبر سلسلة من التحويلات المرتبة والخاضعة لمختلف الشروط إلى البنية السطحية، بينما التحليل الآلي للجملة كان يتبع الوجهة المقابلة بحيث يقوم بالتحويلات في عكس الاتجاه إذ يحلل الجملة من أجل إيجاد البنية التحتية. باستثناء أنحاء المكونات أو الأنحاء المركبية، ليس لتلك البحوث اللسانية المتقدمة صلة بالتقنيات المعروفة في البرمجة.

بالإضافة إلى ذلك، برهنت نتائج **بيترز وريتشي Peters & Ritchie** على أن الأنحاء التحويلية المقترحة في النظرية المعيارية ضعيفة التلاؤم مع الأنحاء الخالية من القيود، بتعبير آخر: إن تلك النتائج أثبتت ضعف الأنحاء التحويلية في توليد اللغات[2]، لقد كان ذلك سببا أدى بعض الحاسوبيين إلى تغيير النموذج النظري للأنحاء التحويلية ببرمجتها على شكل شبكات الانتقال المتصاعد[3].

بالموازاة مع ذلك، تطور أسلوب آخر في البرمجة سمي بالأسلوب **الإثباتي *déclaratif*** أو **البرمجة المنطقية** حيث توصف الخصائص التي يجب أن يتسم بها موضوع ما بمعزل عن

---

[1] المقصود بالتشغيل الآلي هو "إنشاء برنامج حاسوبي يتناسب مع مهمة معطاة مصورنة سلفا. فصورنة النحو لها فائدة للمعالجة الآلية للغة بحيث تسمح ببلورة وتوقع الحساب الذي يفرعه استعمال هذه الأنحاء في تطبيقات دقيقة. يتكلم الباحثون عن تشغيل لنموذج لساني، عندما يعملون على إنشاء مجموعة من البرمجيات التي تسمح بترجمة الأنحاء والمعاجم في صيغة تتناسب مع ذلك النموذج، وتمكّن من توفير تحليل لساني للجمل المعطاة". ينظر: Abeillé (Anne), Les nouvelles syntaxes aaaammaaaaddduccccciiccccaaaaalyeeaaaaFaaççaAAAAAAAAooooiiooooa11111111111111111

[2] ينظر: Trask (R.L.), A dictionary of grammatical terms in linguistics, Graham McFee, 1996, « Peters & Ritchie results ».

[3] شبكات الانتقال réseaux de transition هي تمثيل بياني لإوالية automate تمثل الحالات فيه عند الاقتضاء عبر دوار مرقمة (تسمى أيضا عُقدا أو قِمما)، وكل انتقال يمثل بواسطة سهم معنون x ينتقل من الحالة ei إلى الحالة ej. ينظر: Abeillé (Anne), Les nouvelles syntaxes gggam me eii iii i iiiiii iiiii ii ii iiiii ii i i iiii*op.cit*, p. 302

الإجراءات المستعملة في بناء هذا الموضوع أو في فحص خصائصه. هذا النوع من البرامج يفصل بطريقة واضحة بين المعارف اللسانية وبين الوظائف الحاسوبية، وبالتالي هو سهل التغيير، وسهل تبديل الاتجاه؛ فملفات المعطيات اللسانية يمكن أن تستعمل في التحليل أو التوليد في كلا الاتجاهين[1].

هذه المقاربة تم عرضها بابتكار **لغة برولوغ *Prolog***، وهي لغة تستخدم الرموز لتمثيل العمليات المنطقية وتهدف إلى جعل البرمجة عملية منطقية بحتة، يكثر استخدام هذه اللغة في الذكاء الاصطناعي وتسمى الوحدة الأساسية فيها بالذرة، وتعبِّر الذرة عن علاقة بسيطة بين مفردات محددة[2]. وقد أُعدت هاته اللغة خصيصا لإجراء حسابات على المعطيات المنضدة في مشجرات، ويتم الحساب تصاعديا بتراكم القيود التي تفرض على المعادلات بين المشجرات. في سنوات 1990، طُورت هاته اللغة إذ أدخلت عليها قيود أخرى تتعلق بالأرقام، المجالات، المجموعات المنتهية لذلك تسمى البرمجة باستخدام هذه اللغة **البرمجة بالقيود *contraintes***[3].

وعليه ومن منظور تلك المقاربة المنطقية، لا يعطي النحو قواعد إنتاج الجمل، ولكن مجموعة من القيود التي يجب أن تتوفر في سلسلة حتى تعد جملة صحيحة التكوين، هاته القيود ليست بحاجة لأن تكون مرتبة، على عكس التحويلات التي تتم بالمقاربة الإثباتية، لا نحتاج إلى

---

[1] ينظر: 19- eee ll((((nnn e ee aaaaammaaaaddfffff fff ii Laiiiii eiiiiii iiiiii 0iiiiiiii

[2] ينظر: معجم الحاسبات، مرجع سابق، ص. 116-144.

[3] ينظر: Monin (Jean-François), « Programmation », Encyclopædia Universalis, 2002

عدة مراحل وسطى للتفريع، وإنما يمكن الاقتصار على عمليات تضيف تخصيصات لسانية من دون أن تحذف معلومات وسطى[1].

بالمقابل، طورت الأنحاء التوحيدية من خلال أدوات صورية في مجالات متعددة، خاصة في المنطق والذكاء الاصطناعي، وكذا من خلال إنشاء أدوات برمجية للترجمة الآلية والتصحيح الإملائي وغيرهما ضمن المعالجة الآلية للغة، كل ذلك ساهم بشكل كبير في تطوير نماذج تركيبية جديدة. كما شجع على تطورها تيار حاسوبي يرفض أية نظرية لسانية لتبقى دراسته محصورة في الانتظامات الملاحظة مباشرة في صورة علاقات كمية ضمن مقاربات احتمالية.

وبالتالي يمكن اعتبار الأنحاء التوحيدية أو الأنحاء ذات القيود أساسا نظريات تركيبية جديدة في سنوات 1980، بالإضافة إلى مدرسة كمبردج التي تطورت هي أيضا بعد أن تخلت عن جهازها التحويلي إلى مبادئ أكثر تعميما حول انتقال المكونات. يتعلق الأمر بنماذج تبحث عن مفصلة ظاهرة بين المعجم والتركيب والدلالة، والخصائص اللسانية الممثلة هي عبارة عن معلومات موصولة بالمورفيمات، أو بالمركبات، أو البناءات، مؤلفة بواسطة عمليات متنوعة أهمها عملية التوحيد[2].

إن هذا التصور هو مكسب للمعالجة الآلية للغات لأنه ينطلق من وصف منظومات اللغة في حالة ترابطها حتى تكشف عن مختلف العمليات التي تحدث من أجل تحويل المعطيات

---

[1] ينظر: Abeillé (Anne), Les nouvelles syntaxes rr ammarre ’ iiiii i i i ii ii iii ii ii ii iii ii ii i i i ii , *op.cit*, p. 13.

[2] **توحيد *unification*** بنيتي سمات أ **وب**، الذي يرمز له بـــ **(أ U ب)**، هو البنية الصغرى والامتداد لـــ أ **وب**، تحتوي هذه البنية الموحدة على كل المعلومات التي تحملها البنيتان. ينظر: Bouillon (Pierrette) et *al*., Traitement automatique des langues naturelles, Aupelf-Uref-Ed. Duculot, Belgique, 1998, p. 94.

اللغوية من شكل لآخر في اتجاه الفهم أو الإنتاج اللغويين. وجه آخر لأفضلية تلك النماذج التوحيدية في نظر **أبيه Abeillé** هو اعتمادها على النماذج المنطقية أو الرياضية كأنحاء المكونات، وبنية السمات[1] التي درست وبلورت كمناهج للبرمجة. وعليه، فإن هذه الأنحاء هي على العموم نتيجة لتسوية أو توافق بين تعبيرية لسانية تتمثل في الرغبة في تيسير التعبير عن مختلف المبادئ اللسانية بإضافة متغيرات رمزية، وبين فعالية تتمثل في ترميز القيود والتقليل من العمليات[2].

- ***النمذجة الرياضية:***

بالنسبة للباحثين الذين يميلون إلى المقاربات الرياضية، تمثل النماذج الاحتمالية التي تعتمد على مفاهيم إحصائية أبرز الوسائل الرياضية في حساب الاحتمالات الخاصة بتوالي الكلمات في النصوص. تتميز النماذج الإحصائية بأنها تعطي معلومات كمية حول صحة فرضية، ولكنها بالمقابل تحتاج إلى كمية معتبرة من المعطيات اللازمة لإنشائها[3]. يتعلق الأمر بالملاحظة المباشرة للملفوظات عند دخولها وخروجها في برمجية المعالجة الآلية للغات حيث احتمال حالة مثل: صوت أو كلمة أو ملفوظ تحدده الحالات السابقة، ولا يعتبر فيها تمثيل المعنى ضروريا في الكثير من تطبيقاتها.

---

[1] **بنية السمات *structure de traits*** هي مخطط بأسهم موجهة تحمل عقدا تمثل مختلف السمات اللسانية. تستعمل هذه البنيات من أجل تمثيل المعارف اللغوية في اللسانيات النظرية والحاسوبية أيضا. ينظر: Ligozat (Gérard), Représentation des connaissances et linguistique, Armand Colin, Paris, 1994, p. 5.

[2] ينظر: Abeillé (Anne), Les nouvelles syntaxes rr ammarre ’ iiiii i i iinn n aaaaeeeuuuuaaaaa iii *op.cit*, p. 19.

[3] ينظر: Yousfi (Abdellah),Traitement automatique de la langue (texte et parole), *op.cit*, p.56.

لقد طُورت بنجاح مقاربات إحصائية للتعرف على المنطوق وفي التحليل الصرفي-التركيبي وفي ترجمة النصوص من لغة إلى أخرى، أين يُستغل نموذج **سلاسل ماركوف *de chaînes Markov***، في عملية المعالجة، وقد طوِّر هذا النموذج تحت اسم ***n*-grammes**، وهي سلسلة جزئية مكونة من ***n*** عنصرا مبنية انطلاقا من سلسلة معطاة بحسب احتمال ***n*** حالة سابقة. جاءت الفكرة من أعمال **شانون Shannon** في نظرية الاتصال: تتضمن هذه الفكرة أنه انطلاقا من سلسلة معطاة من الحروف يمكن التكهن بالحرف الموالي. من السهل إذن إنشاء توزيع احتمالي للحرف الموالي لسلسلة سابقة طولها n. هذه النمذجة تتوافق مع نموذج ماركوف من الرتبة n حيث إن n من الحروف السابقة الملاحظة مستعملة للتكهن بالحرف الموالي[1].

تستعمل المناهج الإحصائية للمعالجة حيث يتم اختيار النتيجة الأكثر توقعا إذا تعددت الفرضيات، كما تستعمل لإنشاء الموارد اللسانية كصياغة معجم ثنائي اللغة بملاحظة التماثلات بين النصوص المتوازية في اللغتين، وصياغة معجم أحادي اللغة بحسب السياقات الملاحظة لكل كلمة في المدونة، واستقراء قواعد النحو بحسب متتاليات المقولات الملاحظة في المدونة ووفقا لتواترها. تسمح هذه المناهج بمعالجات فعالة، متينة ومحددة إذ توجد دائما حلا يزيد الاحتمالات إلى أقصى درجة. الفرق بينها وبين المناهج اللسانية أنها توفر موارد لسانية بتكلفة أقل لكنها لا تتمتع بثقة كافية في الاعتماد عليها كتلك التي يوفرها اللسانيون. بالإضافة إلى أنه لا تمكن الاستعانة بها إلا مع وجود مدونات ضخمة محللة سلفا، وأن حساب الاحتمالات يتغير في كل مرة بحسب نوع النص أو ميدان التطبيق.

---

[1] ينظر: ee llrrr (I. Itt dddccii a AAL e à ggggggi egggiiggtttt U sité Lille3, p.20, www.grappa.univ-lille3.fr/polys/info_ling/index.html (تاريخ التصفح: 2010/09/11).

نوع آخر من المقاربات الرياضية هو إواليات الحالات المتناهية، وهي إواليات تتعرف على اللغات التي تولدها الأنحاء الصورية[1]. وهي تشترك مع المقاربات السابقة في كونها لا تسمح إلا بمعالجات محلية لعناصر الرسالة التي تحمل معلومات في سياق منته. بالمقابل، يسمح هذا النوع بتمثيل أكثر تعقيدا للمعارف اللسانية، كما أنه يضمن معالجات سريعة ومحددة كثيرة الفائدة في التحليل المعجمي أو التركيبي، وفي استخراج المعلومات من نصوص كبيرة الأحجام.

من أجل الحصول على أفضل أداءات المعالجة، يجتهد الباحثون في تأليف عدة مناهج خاصة المقاربات اللسانية والرياضية.

- ***المقاربات المعرفية***

في هذه الاستراتيجية، تندرج المعالجة الآلية للغات ضمن ميدان العلوم المعرفية[2] حيث يتعلق الأمر بنمذجة الفرضيات حول اشتغال اللغة الإنسانية واختبارها في الآلة، وإن لم يتم الحصول على أفضل التطبيقات العملية.

بعض المحللات، ذات الصلة بعلم النفس اللساني، عوض أن تقوم باستقراء كل الفرضيات الممكنة بطريقة مملة، تحاول أن تركز على الكلمات الأكثر أهمية كما يفعله الإنسان عادة خلال قراءته السريعة. كما أن بعض البرمجيات تقوم بالتعرف الخطي على كلمات المدونة من خلال التركيز مباشرة على الأحياز الأقرب إلى الحالة المثالية التي زودت بها ذاكرة البرمجيات، ثم تعمل على التعرف على صور تلك الكلمات انطلاقا من موقعها في السياق.

---

[1] ينظر: Ligozat (Gérard), Représentation des connaissances et linguistique, *op.cit*, p. 39

[2] العلوم المعرفية هي جملة من العلوم تدرس اشتغال الذهن والذكاء وتعنى بنمذجته، ودراسته مبنية على تظافر مجموعة من الاختصاصات (الفلسفة، علم النفس، الذكاء الاصطناعي، علوم الأعصاب، اللسانيات، الأنتروبولوجيا). ينظر: الزناد (الأزهر)، نظريات لسانية عرفنية، منشورات الاختلاف، الجزائر، ط.1، 2010، ص. 15.

ضمن هذه المقاربة أيضا، وبالصلة مع مبحث الجهاز العصبي، طورت في نهاية سنوات 1980 شبكات عصبونات اصطناعية، وهي معمارية لإواليات أساسية تشتغل مثل خلايا شديدة الترابط فيما بينها، استخدمت بداية في التعرف على الأشكال، ثم في ميدان المعالجة الآلية للغات، وهي قطاع جديد يتلاءم جيدا مع التعلم الذاتي ومع تمثيل الظواهر المستمرة مثل التغيرات الصوتية أو تحديد المعاني، ولا تحتاج عملية المعالجة إلى مدونات ضخمة مثل المقاربات السابقة[1].

4.2. موارد المعالجة

في ميدان المعالجة الآلية للغات، يحتاج الباحثون إلى مجموعة من الموارد المهمة، يتعلق بعضها بوسائل حاسوبية التي تمنح البرامج القدرة على الحساب، ويتعلق بعضها الآخر بقواعد البيانات[2] المزودة بشتى المعارف اللسانية.

- ***الموارد اللسانية***

تشكل المعاجم والمراجع النحوية موارد تقليدية ومهمة للتزويد بمختلف المعارف المتصلة باللغات، لكنها قليلة الاستعمال في ميدان المعالجة الآلية لعدة أسباب منها: أن تلك الموارد التقليدية مهما كبر حجمها فهي ناقصة وغير مصورنة؛ إذ أن أفضل المعاجم لا تزود إلا بالمفردات في شكلها القاعدي، وتهمل بالتالي أسماء الأعلام والمفردات المولدة والمشتقة الجديدة؛ كما أنها تقدم معاني المفردات في شكل غير صوري يتضمن نصوصا حرة وتعاريف

---

[1] ينظر: Abeillé (Anne), « Traitement automatique des langues » *in* Encyclopædia Universalis, Paris, 2002.

[2] قواعد البيانات مجموعة بيانات مسجلة في ملفات على نحو يحدد الروابط المنطقية بين نوعياتها المختلفة. ينظر: معجم الحاسبات، مرجع سابق، ص. 53.

وأمثلة، والمعلومات المعجمية والنحوية التي تحملها المفردات ليست مترابطة فيما بينها ونادرا ما تبين المعلومات النحوية في المعاجم. إذن، صار من الضروري إنشاء معاجم إلكترونية بأحجام ضخمة مهيئة لاحتياجات المعالجة الآلية للغات، وقد حقق جزء منها فعلا بالنسبة لأبرز اللغات الأوروبية، يمكن أن نذكر مثلا **القاموس التفسيري التأليفي للفرنسية المعاصرة** ***Dictionnaire explicatif et combinatoire du français contemporain*** الذي يتميز بأنه تفسيري حيث يصاحب كل عنصر معجمي بتفسير صوري لمدلوله، ووضع تمثيل دلالي له يواكب شرحه المعجمي، كما أن سمته التأليفية تجعله يعطي قيمة خاصة للتأليف المعجمي أي لتسلسل الوحدات المعجمية في النص[1].

وأما بالنسبة للمعارف المتصلة بأصوات اللغة، فقد توصل الباحثون إلى إنشاء قواعد بيانات صوتية تخزن المقاطع الممكنة، وطريق نسخ الكلمات. وبالنسبة للمعارف النحوية، فقد تقدم الحديث عن بعض النماذج المستخدمة في المعالجة الآلية للغات كالأنحاء الصورية، لكن القليل من الأنحاء الإلكترونية تم تحقيقها إلى اليوم.

- ***الموارد الحاسوبية***

تحتاج المعالجة الآلية للغات إلى موارد حاسوبية متطورة مثل كل القطاعات التي تعمل على حوسبة مشكلاتها المعقدة. بعض البرمجيات تشتغل في الحاسوب، لكن أغلب البرمجيات الذكية منها تستلزم آلات أكثر تقنية من الحاسوب الشخصي؛ إذ لم يمكن الحصول على آلات ذات قدرة كافية على معالجة أحجام ضخمة من النصوص باستخدام معاجم إلكترونية في وقت

---

[1] ينظر: ’’’’ č (Ioo)))))*al.*, Introduction à la lexicologie explicative et combinatoire, Duculot, Belgique, 1995, p. 10.

معقول إلا حديثا مع تطور تكنولوجيا الحاسوب. كما أن الشاشات ولوحة المفاتيح قد كيفت مع أنماط الكتابة والعرض لتشمل الحروف غير اللاتينية والكتابات غير الألفبائية. بالإضافة إلى ذلك، سمح التطور في ميدان الإلكترونيات أيضا بتحسين الأدوات الصوتية كظهور البطاقات الصوتية لتأليف الكلام، وكذا التسجيل الرقمي للتعرف على المنطوق.

مثل كل المهام في الذكاء الاصطناعي، تطرح عدة مشاكل من بينها لغة البرمجة، والمعمارية وهي الشكل العام لنظام الحاسوب، وتتضمن أيضا مجموعة التعليمات في الحاسوب وطريقة تعامل المستخدم معه وتنظيم الذاكرة وطرق العنونة وعمليات الإدخال والإخراج والتحكم، والخوارزمية التي هي مجموعة محددة من خطوات منطقية وحسابية تحدد المنهاج لحل مسألة ما[1].

لغات البرمجة 'العالية' التي هي بعيدة جدا عن التمثيلات الثنائية الداخلية للآلة مثل **لغة برولوغ *Prolog***، تستخدم بكثرة في المعالجة الآلية للغات، وقد ابتكرت من أجل سد حاجياتها، هذه اللغة تتضمن قدرات الاستدلال الخاصة باللغات المنطقية، وتتلاءم جيدا مع كتابة البرمجيات الصغيرة. تستعمل أيضا **اللغات الوسطى**[2] ***orientés objets*** التي تتضمن أنظمة طبيعية للتصنيف، كما صممت حديثا لغات ***Perl*** و***Awak*** لمعالجة الوحدات النصية.

---

[1] ينظر: معجم الحاسبات، مرجع سابق، ص.11، 7.

[2] اللغات الوسطى orientées objets هي لغة يحول المترجم إليها برنامجا مكتوبا بإحدى لغات البرمجة، وهي كثيرة الاستعمال في الذكاء الاصطناعي ومكيفة مع تمثيل المعارف ظهرت في سنوات 7970. ينظر: Stehlé (Jean-Luc) et Hochard (Pierre), Ordinateurs et langages, Ellipses, Paris, 1989, p. 117.

فيما يخص المعمارية أو بنية الحاسوب، تضع تقنية الإلكترونيات الصغرية بين أيدي المستخدمين آليات لمعالجة المعلومة فائقة السرعة حيث الانتقال يتم في أقل من جزء واحد من المليار من الثانية، كما أن أجهزة التخزين في الحاسوب تطورت بحجم وسرعة متعاكسين في الاتجاه، فكلما زاد مكون من مكونات التخزين سرعة زاد حجمه صغرا.

وبما أن المطلوب هو توفير أجهزة ذات قدرة تخزين كبيرة وسرعة فائقة في الآن نفسه، فإنه يجري اتباع استراتيجية معينة تتمثل في استعمال ذاكرة سريعة ذات حجم صغير لتخزين معلومات مفيدة في لحظة معينة بالاعتماد على مبدأ **الموضعة *localité*** التي تتميز بها المعلومة، والتي تجعل احتمال إعادة استعمال المعلومة ذاتها يفوق بكثير احتمال الحاجة إلى استخراج غيرها[1].

5.2. أدوات المعالجة

إذا أردنا أن ننجز نظاما أو برنامجا للمعالجة الآلية للغات، فما هي المقاربة التي يمكن أن نوظفها في تحقيقه ؟

هناك مقاربتان يمكن الاستعانة بهما لتحقيق نظام في مجال المعالجة الآلية للغات: المقاربة المباشرة وتتضمن ابتكار التطبيق باستخدام لغة برمجة مثل: ***Lisp, Pascal, Prolog, C...*** من أجل كتابة برنامج يمكن أن يؤدي الخدمات المستهدفة، وهذه هي المقاربة التي نختارها من أجل القيام بتوليد الجمل في اللسان العربي معتمدا على لغة برولوغ ***Prolog***. والمقاربة الثانية وهي الأنسب بحوسبة المعطيات اللغوية، وتتضمن استخدام أدوات حاسوبية متنوعة

---

[1] ينظر: أنسو (فرانسوا)، « من **فونيومن Von Neumann** إلى المعالجات الصغرية الفائقة »، ما التكنولوجيا ؟ تر. محمد نايت الحاج وعبد الهادي إدريسي، المجلس الأعلى للثقافة، القاهرة، 2005، ج. 5، ص. 145-747.

مبتكرة خصيصا للأوصاف اللسانية بهدف ترميزها في قالب يسير التناول بالنسبة للساني، وفي الوقت نفسه يسهل على الحاسوب ترجمتها آليا إلى برنامج.

الأدوات الحاسوبية من أجل المعالجة الآلية للغات تستخدم في عدة سياقات ولأهداف مختلفة، من أهمها مجال التعليم من خلال تكوين الباحثين، وتوضيح مكنزمات المعالجة الآلية للغات، والبرهنة على اشتغال النظريات النحوية، وكذلك تستخدم في مجال البحث كدراسة تقنيات جديدة حاسوبية ووصفية، وإقامة أوصاف لسانية متطورة للمعالجة، وإنتاج مختلف التطبيقات اللسانية الحاسوبية[1].

3. محاور البحث في "المعالجة الآلية للغة"

يصنف بعض الباحثين ميادين البحث في المعالجة الآلية للغة الطبيعية إلى أربعة محاور بحثية هي الهندسة اللغوية، واللسانيات المعلومية، والمعلوميات اللسانية، واللسانيات الحاسوبية[2].

المحور الأول: الهندسة اللغوية

وتهتم بإيجاد الطرق التي تساعد على تنظيم المعطيات اللسانية وتشغيلها وتخزينها، وذلك بهدف إنشاء برامج قادرة على أن تقوم بمعالجة خاصة لأحد جوانب اللغة بفعالية. هذا المحور متصل مباشرة بالمعلوميات ومرتبط بالمتطلبات المادية للحواسيب الآلية. ينتظر من الباحثين في هذا المجال إيجاد أفضل الحلول الخوارزمية، وإنجاز بنى للمعلومات داخل الآلة. يشكل هذا المجال البحثي سمة الاستعمال أو التطبيق حيث يوظف المهندس نتائج اللسانيات إذا أمكن أن

---

[1] ينظر: Bouillon (Pierrette) et *al*., Traitement automatique des langues naturelles, *op.cit*, pp. 201-202.

[2] ينظر: Ratté (Sylvie), *Interprétations des structures syntaxiques : une analyse computationnelle de la structures des évènements*, *op.cit*, pp. 50-51.

تساعد في التطبيقات المعلومية دون أن تشارك اللسانيات النظرية مباشرة في هذه العملية، مما يعني أن هناك عدم تخطيط لاستثمار النتائج اللسانية في التشغيل المعلومي.

المحور الثاني: اللسانيات المعلومية

المجال الموالي الذي يسمح بمعالجة اللغة بطريقة آلية هو اللسانيات المعلومية، ولكن هذه المرة تتم المعالجة من وجهة نظر الباحث اللغوي الذي ليست له بالضرورة رغبة في إنجاز برامج ذات فعالية وصرامة معلومية بقدر ما يرغب في اختبار التصورات النظرية المتعلقة باللغة حتى تحدث تلك البرامج الحاسوبية المنجزة ردة فعل قد تسهم في تعديل تلك التصورات النظرية، بعبارة أخرى: تشكل البرامج الحاسوبية التي يبغي إنشاءها مختبره العلمي الذي يجري فيه تحاليله وتجاربه على تصوراته حول اللغة.

إن التشغيل المعلومي في هذه الحالة يفيد النظرية اللسانية ويعطيها إيضاحات صورية، ويدفع الباحثين إلى إعادة النظر في منجزاتهم العلمية وفي مكونات تلك النظرية حتى تستجيب للمعالجة، لا يعني ذلك تطويع النظرية اللغوية حتى تلائم متطلبات الحاسوب، ولكن البرامج الحاسوبية المنجزة قد تكون مجالا مهما وفعالا لتجريب النظريات اللسانية. يأمل الباحثون في هذا المجال أن يقيموا جسر تواصل بين النظرية اللسانية والتشغيل الآلي، لذلك يشكل هذا المجال البحثي بدوره سمة التكييف.

المحور الثالث: المعلوميات اللسانية

المحور الثالث من محاور البحث في ميدان المعالجة الآلية للغة الطبيعية والذي يمثل فصيلا هجينا من المحورين الأولين، نجده خاصة في أوساط الذكاء الاصطناعي، لذلك يهتم

الباحثون فيه بتقييس (محاكاة) سلوك لساني معين في الحاسوب، ويتم هذا إما بتكييف مستويات التمثيل المقترحة مع إحدى النظريات اللسانية، أو بإقامة نظرية لتلك المستويات حيث تكون التفاعلات ضرورية بين التصورات اللسانية والوسائل الحاسوبية من وجهتي النظر العملية والوظيفية: العملية ونقصد بها التفاعلات التي تخص إجراءات الانتقال من مستوى لساني لآخر، أما الوظيفية فنقصد بها التفاعلات التي تتعلق بالأدوار التواصلية التي تؤديها.

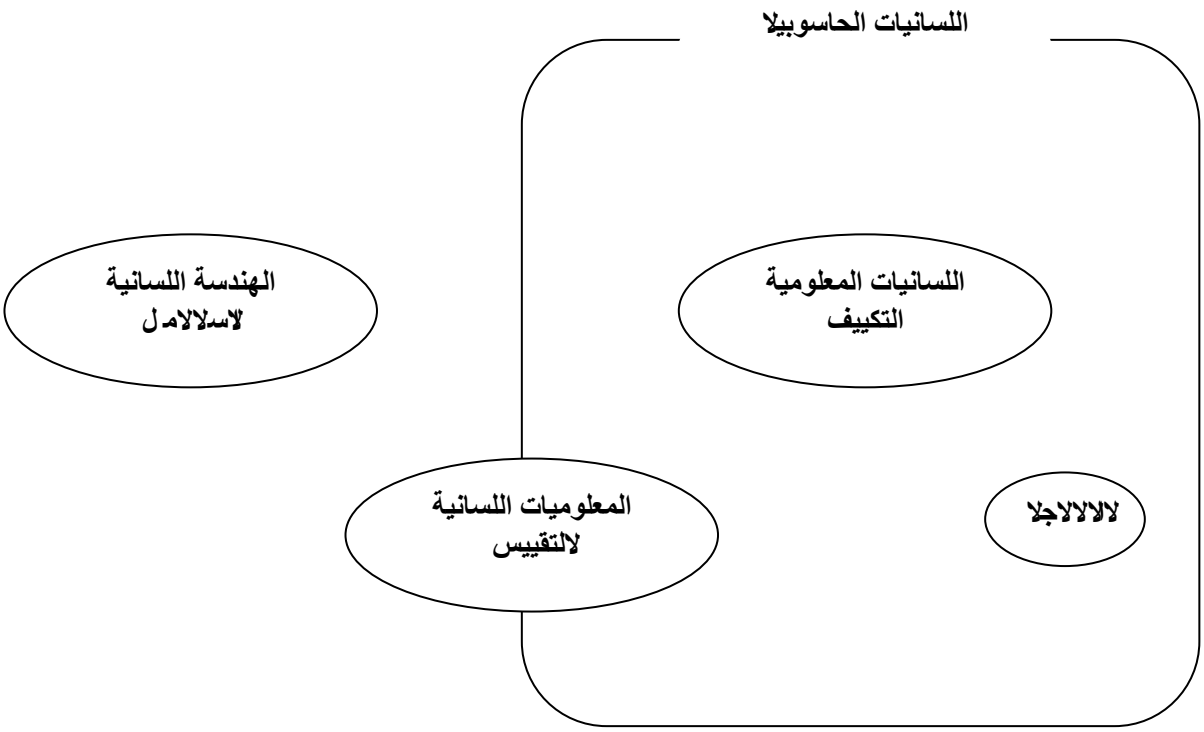
يمكن لمستويات التمثيل أن تتطابق مع المستويات التي تقترحها النظرية اللسانية، مثل التعرف على الجمل المكتوبة يستدعي وجود منظومة للتعرف على الحروف، ثم أخرى للتعرف على الكلمات، وأخيرا يضم منظومة للتعرف على الجمل، وهذا الذي تبينه النظرية اللسانية من وجود ثلاث مستويات للتحليل: فونولوجي، صرفي، تركيبي، يضم كل مستوى منها وحدات لغوية معينة. يتغيا الباحثون في هذا المجال الحصول على سلوك لساني يكافئ ويعادل ذلك الذي يشاهد لدى الإنسان. يمثل هذا المحور سمة التقييس.

المحور الرابع: اللسانيات الحاسوبية

لا تشكل اللسانيات الحاسوبية محورا مستقلا متميزا، ولكنها بالأحرى تجمع المحورين الأخيرين بإضافة سمة ثالثة للبحث هي سمة النمذجة. يهتم الباحث هنا بإقامة فرضيات حول مستويات التمثيل الضرورية لتفسير مظهر حاسوبي للغة، وتطوير نماذج نظرية لتفاعلات تلك المستويات. وبالتالي، فإن هذا المحور البحثي لا يهدف إلى تشغيل نظرية لسانية، كما هي الحال في اللسانيات المعلومية، كما أنه لا يهدف إلى تقييس سلوك لساني خاص في الحاسوب، كما هي الحال في المعلوميات اللسانية، لكنه يقترح نماذج حاسوبية لبعض النواحي الحاسوبية في اللغة؛

إنه يقترح تحليلا للعمليات القاعدية، ولمستويات التمثيل والتفاعلات القاعدية التي تجري لدى إنتاج اللغة أو تحليلها. أخيرا، إن هذه النماذج لا تدعي إعادة إنتاج سلوك لساني خاص، ولا تحقيق اشتغال نظرية لسانية معينة، ولكنها تضطلع بتفسير الجانب الحاسوبي للغة اعتمادا على أدوات رياضية (آلات صورية، تحويلات جبرية...).

الشكل التالي يعكس تصنيف ميدان المعالجة الآلية للغة الطبيعية إلى تلك المحاور الأربعة المقدمة سلفا[1].



الشكل8: محاور البحث في المعالجة الآلية للغة

[1] ينظر: Ratté (Sylvie), *Interprétations des structures syntaxiques : une analyse computationnelle de la structures des évènements*, *op.cit*, p. 52.

يبين هذا الشكل تصنيف محاور البحث في ميدان المعالجة الآلية إلى أربعة محاور يغلب على كل واحد منها أسلوب في التناول الآلي للظاهرة اللسانية وفي الهدف من هذا التناول، أما الهندسة اللسانية فتقوم على استعمال المعلومات اللغوية لأغراض حاسوبية وتطبيقها في مختلف البرامج المنجزة، وأما اللسانيات المعلومية، فتعتبر مختبرا حاسوبيا لفحص مختلف النظريات اللسانية وتكييفها مع الآلة. والمعلوميات اللسانية مجال لإنشاء البرامج الحاسوبية التي تعمل على تقييس النشاط اللغوي الإنساني. وأخيرا، محور اللسانيات الحاسوبية الذي يضم محوري اللسانيات المعلومية والمعلوميات اللسانية إذ يتبع الباحث فيه أسلوبي التكييف والتقييس ويضيف إليهما أسلوبا آخر هو أسلوب النمذجة في تناوله للجوانب الحاسوبية للغة والتي تتعلق بأنظمة بناء الجملة وعناصرها المتعددة في شتى المستويات.

إن اللسانيات الحاسوبية ومختلف المجالات التطبيقية للمعالجة الآلية للغات الطبيعية التي تتناول توليد الجمل تنطلق من مجموعة من الأوصاف والنماذج النظرية التي توفرها اللسانيات واللسانيات النفسية، من أجل ذلك أحاول أن أتعرض في الفصل الموالي إلى أبرز النماذج الخاصة بتوليد الجمل، والتي عمل اللسانيون واللسانيون النفسانيون على صياغتها وتشغيلها.

الفصل الثالث:

نماذج التوليد اللسانية واللسانية النفسية

الفصل الثالث: نماذج التوليد اللسانية واللسانية النفسية

إن توليد الجمل هو ظاهرة لسانية على اعتبار ما تؤول إليه هذه العملية التي تتفرع عنها مجموعة من الجمل ذات طبيعة لغوية إذ هي مكونة من ألفاظ تحمل دلالات معجمية وصرفية وتركيبية، وبالتالي يمكن دراسة نظامها واقتراح النظريات اللسانية في وصفها. كما يمكن اعتبارها ظاهرة لسانية نفسية من وجهة النظر الحركية والدينامية التي تتم فيها عملية التوليد بحيث يقع الربط بين معالجات متتابعة ومترابطة تتناول الكيانات اللغوية وذلك من المستوى الأكثر تجريدا والأقرب إلى الدلالات إلى المستوى الأقل تجريدا والأقرب إلى الألفاظ، وعليه يقترح الباحثون نماذج يتصورون من خلالها الكيفية التي تم بها إجراء هذه العملية. هذا الذي دفعني إلى تخصيص فصل كامل من البحث أتعرض فيه إلى مفهوم توليد الجمل بين اللسانيات واللسانيات النفسية.

بداية، يعرف المتكلمون والمناطقة التوليد بأنه: إيجاب شيء لشيء، أو استلزامه له، أما معناه الحسابي –والمستعمل عند الرياضيين المحدثين– فيدل على التوسل بالطريق التكراري في إنتاج المطالب. وقد استثمر المناطقة هذا المفهوم في صنع لغاتهم الصورية، واللسانيون أيضا في وضع أنحاء اللغات الطبيعية. ويقوم مجمل الطريق الذي يتوسل به في إنتاج المطالب على ثلاث قواعد: القاعدة الأولى ابتدائية تقتضي تحديد المجموعة الأصلية من المطالب، والقاعدة الثانية تكرارية تقتضي تفريع المطالب الأخرى بواسطة تطبيق هذه القاعدة على المطالب المتقدمة عددا متناهيا من المرات، والقاعدة الثالثة ختامية توقف تحصيل المطالب على التوسل بالقاعدتين السابقتين: الابتدائية والتكرارية.

ويوضح **طه عبد الرحمن** ذلك بمثال مجموعة الأعداد الطبيعية ط؛ إذ بالإمكان توليد ط بواسطة القاعدة التالية: "أضف 1"، وذلك بأن يضمن المجموعة الأصلية عنصرا واحدا هو o فيكون بذلك عددا طبيعيا، ثم يفرع عليه أي عدد طبيعي آخر بواسطة القاعدة المذكورة مطبقة عليه وعلى الأعداد التالية له واحدا واحدا حتى الوصول إلى العدد المرغوب فيه، وبالتالي فإن القاعدة "أضف 1" هي عبارة عن طريق تكرارية لتوليد العدد المطلوب[1].

بعد تعريف التوليد، أحاول أن أتعرض فيما يأتي إلى النماذج اللسانية وكذلك إلى النماذج اللسانية النفسية التي تناولت موضوع توليد الجمل.

1. النماذج اللسانية

إن النظريات اللسانية المختلفة نادرا ما تقدم نفسها بشكل واضح على أنها نماذج لإنتاج الكلام وتوليد الجمل. قد يرجع السبب في ذلك إلى أن البحث في سيرورات إنتاج الكلام وفهمه من مهام علم النفس اللساني، وليست من مهام اللسانيات التي ينحصر حقلها في دراسة بنية اللغة وانتظام مكوناتها في مختلف المستويات دون العروج على الميكنزمات التي تؤدي إلى معالجة المعلومة اللغوية إنتاجا وفهما وتحويلا وتخزينا. ومع ذلك، يعتمد الباحثون على بعض تلك النظريات اللسانية كأساس نظري من أجل إنشاء نماذج الإنتاج، خاصة تلك النظريات التي تشارك العلوم المعرفية اهتمامها بالملكات المعرفية عند الإنسان والتي تشكل اللغة أبرزها.

وبالتالي، يعد الباحثون من تلك النظريات اللسانية المؤثرة في مجال إنتاج الكلام أو توليده تيارين لسانيين. يختلف ذانك التياران من حيث مبادؤهما المتصلة بطبيعة اللغة، وكذا من

---

[1] عبد الرحمن (طه)، اللسان والميزان، المركز الثقافي العربي، الدار البيضاء، ط.1، 1995، ص. 27.

حيث المناهج التي تترتب عن تلك المبادئ. التيار اللساني الأول هو صوري ممثله النحو التوليدي خاصة، والتيار اللساني الثاني هو وظيفي أو حركي تمثله عدة نظريات وظيفية ومعرفية[1].

من مبادئ التيار الصوري أن الصورنة ليست فقط وسيلة إجرائية للتحليل والتنظير، فقد تتقاسم معها اتجاهات أخرى هذه الرغبة كعلم الدلالة الصوري والمعالجة الآلية للغة، ولكنها تعكس الحالة الأنطلوجية التي تخصَّص للبنى الصورية كما تظهر في التركيب، فهذه البنى الصورية تمثل أساس اللغة التي تعتبر نظاما من القواعد الصورية الثابتة والكلية والفطرية. وعليه، فإن موضوع اللسانيات هو تعيين النحو الكلي، ومنهجها الفرضي–الاستدلالي المبرر بمبدإ فطرية القواعد، يتضمن توليد الجمل الممكنة انطلاقا من هذه القواعد، يشتغل اللساني إذن بالتوليد وبانعزال عن السياق. في الأخير، تعتبر اللغة هنا نظاما مغلقا يمكن وصفه عبر نماذج صورية (أمثلات) من النوع المنطقي–الجبري[2].

بالمقابل، النظريات المتنوعة التي تنتمي إلى التيار الوظيفي أو الحركي تضع فرضية أن البنى اللسانية تحدَّد بواسطة الوظائف التي تؤديها اللغة والتي هي متنوعة. تنوع تلك النظريات نابع من اعتمادها على وظائف متعددة، لكنها متفقة حول وظيفتين عامتين: وظيفة رمزية تتضمن تواضع المتحدثين على استخدام رموز معينة، ووظيفة اتصالية من خلالها يتفاعل أولئك

---

[1] ينظر: Fayol (Michel), Production du langage, Lavoisier, 2002, pp. 67-68

[2] ينظر: Mounin (Georges), « Le modèle en linguistique », Encyclopædia Universalis, Paris, 2002, tome XV, p. 304.

المتحدثون، وبالتالي، لا تمكن دراسة اللغة بمعزل عن الغايات والمقاصد وأوجه استعمالها وسياقاتها (سياق كلامي ومقامي، أدوار المتحدثين، انعكاسات معنى الملفوظات عليهم).

تعتبر اللغة هنا نظاما مفتوحا وقابلا لإعادة التشكل في الإطار الداخلي للغة من خلال تنوع الاستعمالات للأشكال نفسها، أو في الإطار الخارجي للغة من خلال التفاعلات التي تكون بين لغة وأخرى. بالنسبة للساني، يتعلق الأمر بوصف الخصائص الوظيفية للغات انطلاقا من الأشكال المتغيرة واستعمالاتها. لذلك، يفضل الباحثون ضمن هذا الاتجاه المنهج الاستقرائي حيث يبدأ التحليل من الاستعمالات في سياقاتها الطبيعية ويصل إلى القيمة الثابتة للأشكال والمكنزمات. طبيعة القيم الثابتة المبحوث عنها ومستواها هما اللذان يصنفان مختلف المدارس داخل هذا التيار:

الثوابت الطبولوجية ← الطبولوجيا الوظيفية[1]

الثوابت المعرفية ← الأنحاء المعرفية

المبادئ الكلية للتفاعل ← التحليل التحادثي والتداوليات

مكنزمات التلفظ ← نظرية التلفظ...

وتنتمي صورانياتها إلى النوع الموقعي الحركي.

إن التأثير المختلف لهذه النظريات المتنوعة في مجال نماذج الإنتاج يفسَّر من ناحية باختلاف منهجياتها حيث إن النظريات التي تعمل على التعرف تقود إلى نماذج للفهم قد تكون غير صالحة بالضرورة لأن تكون نماذج للإنتاج. النحو التوليدي من بين تلك النظريات له تأثير

---

[1] الطبولوجيا هندسة لا كمية فرع الرياضيات تعنى بدراسة موقع الشيء الهندسي بالنسبة إلى الأشياء الأخرى، لا بالنسبة إلى شكله أو حجمه. إدريس (سهيل)، المنهل، دار الآداب، بيروت، ط.11، 2003، ص. 1208.

معتبر في مجال العلوم المعرفية حيث تعتبره النموذج اللساني المرجعي سواء أصرحت بذلك أم لم تصرح.

كما أن التيار الوظيفي التقليدي بقي بالمقابل لمدة طويلة بعيدا عن حقل العلوم المعرفية لتمسكه المنهجي بوصايا دي سوسير باجتناب السيرورات الذهنية التي يعتمد عليها اشتغال اللغة. ومع أن المبادئ المتعلقة بطبيعة المكونات اللسانية وبطبيعة ميكنزمات المعالجة التي يقوم عليها النحو التوليدي قد تبدو أوضح وأمتن من تلك التي يقوم عليها التيار الحركي إلا أنه قد ظهرت للعيان نماذج لسانية جديدة لم تحافظ على تلك الأسس التوليدية[1].

1.1. الأنحاء التوليدية

لقد أحدثت الأنحاء التوليدية ثورة في اللسانيات الحديثة من خلال المبادئ التي تبنتها، والمناهج التي اعتمدت عليها لدراسة اللغات الطبيعية واللغات الصورية، تلك الأنحاء تأثرت كثيرا بأعمال تشومسكي، لكن بعضها وافقه في بعض تلك الأعمال فقط.

1.1.1. النحو التوليدي: نموذج تشومسكي

يتميز هذا النموذج بأولية التركيب على سائر مكونات النحو، فيما يخص طبيعة اللغة: حيث إن النحو الكلي هو عبارة عن مجموعة من القواعد التركيبية، أو فيما يخص اشتغالية اللغة إنتاجا وفهما. ففي مجال الإنتاج، تعمل البنى التركيبية على توليد الملفوظات أين يتم اختيار الكلمات داخل التركيب الذي هو المكون التوليدي الوحيد، في حين أنه لا مكان لنظرية أو مكون دلالي في هذا النموذج، لأن الدلالة هي نتيجة تأويلية تنبع من إدماج المكونات المعجمية داخل

[1] ينظر: Fayol (Michel), Production du langage, *op.cit*, pp. 68-69

البنية التركيبية. تمثل نظرية العمل والربط آخر ما وصلت إليه النظرية التشومسكية من تطور؛ حيث تهدف إلى إنشاء نحو كلي أي ترمي إلى إنشاء نظرية للخصائص الكلية للغات الطبيعية، وهي نظرية للكفاية اللغوية تأخذ بعين الاعتبار ظواهر اكتساب اللغة[1].

اعتمادا على هذه المبادئ المتنوعة، نبع نموذج تسلسلي أو منظومي لإنتاج الكلام مركزه التركيب، وسيرورات المعالجة في هذا النموذج تتناسب مع إنشاط مختلف المستويات البنوية حسب التراتبية الموضوعة سلفا في النموذج التوليدي. وعليه، فإن النماذج التقليدية للإنتاج ترتكز على هذا النموذج معتقدة في الوقت نفسه تناسبا بين العمليات المعرفية والمكونات اللسانية (الواقع الذهني للنحو الكلي)، واستقلالية لمختلف سيرورات المعالجة التي تتناسب مع المنظومات (الاستقلالية الوظيفية للمكونات اللسانية)، وأولوية للمعالجة التركيبية في عمليتي الإنتاج والفهم (المحافظة على الدور التوليدي للتركيب). حتى في مجال العلوم العصبية، يتجلى أثر تلك المبادئ في النظريات التي تفترض أن الكفاية اللغوية تقع في حيز خاص من الدماغ، وتبحث عن المقابل الفيزيولوجي للعضو الذهني عند **تشومسكي Chomsky**[2].

2.1.1. النماذج التوليدية غير التشومسكية

إن الرؤية التوليدية التي تعطي الأولوية للتركيب وتجعل المكونين الدلالي والصوتي تأويليين نابعة من تشبيه النحو بخوارزم حاسوبي يولد الجمل النحوية، وهي تتزامن مع الرؤية الحاسوبية التي كانت سائدة والتي تعتقد بأن الخوارزمات لا يمكن أن تشغل إلا بشكل تكراري

---

[1] ينظر: Fuchs (C.) et Le Goffic (P.), Les linguistiques contemporaines: repères théoriques, Hachette, Paris, 1992, p. 85.

[2] ينظر: Fayol (Michel), Production du langage, *op.cit*, pp. 70-71

متبعة عددا من الخطوات المرتبة في إنتاج الجمل. هذا التصور أعاد النظر فيه بعض تلاميذ **تشومسكي** منهم **جاكندوف Jackendoff** خاصة بعد أن تبين في الثمانينيات من القرن العشرين ونتيجة لتطور علوم الحاسوب بأن الخوارزمات يمكن أن تشتغل بشكل متواز، ويمكن أن تكون نموذجا مقبولا في تمثيل الاشتغال الذهني[1].

إن هذا النموذج التوليدي التعاقبي المقترح لا يتضمن مكونا توليديا واحدا بل ثلاث مكونات (فنولوجي ودلالي بالإضافة إلى التركيبي) تساهم كلها في بناء الجملة. هذه المكونات هي ذات طبيعة توليدية لأنها محددة بواسطة قواعد تكوينية تسمح بتوليد البنى في كل المستويات، لكنها متعددة ومتصلة فيما بينها بواسطة مكونات تصافحية تعيّن قواعد التوافق أو التجميع بين البنى المستقلة.

في هذه المعمارية الموازية، المحتوى الإعلامي لجملة هو إذن موزع في ثلاث بنيات مستقلة وكذلك الصلات التي تربط بينها. وهكذا فإن العلاقات الدلالية بين الجمل (مثل بناء الفعل للمعلوم وبنائه للمجهول) لا يُنظر إليها على أنها نتاج عملية تحويلية لبنية أولية، ولكن ينظر إليها على أنها طريقة مغايرة للوصل بين البنى التركيبية وبين الأدوار الدلالية.

تتقابل مع مقاربة **تشومسكي** التركيبية التفريعية نظرية للصلات التجميعية. فيما يخص سيرورات المعالجة، يبقى هذا النموذج منظوميا أي إنه يتكون من مكونات مستقلة فيما بينها، لكن الميكنزم الأساسي يكمن في سيرورة توحيدية تسيطر على قواعد التجميع بين مختلف المكونات.

---

[1] ينظر: الزناد (الأزهر)، نظريات لسانية عرفنية، مرجع سابق، ص. 64.

2.1. الأنحاء المعرفية والوظيفية المعرفية

مختلف النظريات التي تتبع التيار الحركي لها تصور مناقض حول طبيعة المكونات اللسانية ودورها. النحو بالخصوص ليس نظاما صوريا خالصا، ولكنه ترتيب خاص لتشفير المعلومة التحادثية (هذه النظرة الوظيفية)، ويحمل هو نفسه معنى باستمرار مع الدلالة المعجمية (هذه نظرة الأنحاء المعرفية). ترتكز هذه المقاربة على تصور خاص للكفاية اللغوية ووظائفها، وتوحِّد هاتين المدرستين اللتين تحاولان سوية إنشاء قواعد جديدة لعلم نفس اللغة[1].

الوظيفية والأنحاء المعرفية تعتبران اللغة ظاهرة تكييفية[2] فيما يتصل بظهورها تاريخيا أو فيما يخص اشتغاليتها. وهذا يتعارض مع المبدأ الفطري لقواعد اللغة عند التوليديين. وعلى نقيضهم أيضا، تعتبر الأنحاء المعرفية اللغة قبل كل شيء نظاما سيميائيا تتحدد بنياته بحسب الوظائف التي يؤديها في إطار الملكات المعرفية العامة. فليست الكفاية اللغوية إذن بمقدرة مكتفية بذاتها، ولا تقبل الوصف بمعزل عن العمليات المعرفية، وتمثل تطورا واستقرارا لأبنية ذهنية معرفية عامة تستوي في نظام لغوي مخصوص أثناء الاكتساب وأثناء توظيفها في التواصل اليومي. وهي في جميع ذلك مرتبطة بعوامل التجربة ومحكومة بظواهر نفسية ليست لغوية في طبيعتها[3]. بسبب رسوٍّ وتجذر اللغة في التجربة، وبسبب المواقف المنهجية الآنفة

---

[1] ينظر: Fayol (Michel), Production du langage, *op.cit*, pp. 71-72

[2] حسب هذا التصور، ابتكر الإنسان اللغة ليتكيف مع واقعه البيولوجي والاجتماعي، وليست اللغة عنده موجودة بالفطرة.

[3] ينظر: الزناد (الأزهر)، نظريات لسانية عرفنية، مرجع سابق، ص. 98-99.

الذكر، تتبنى الأنحاء المعرفية مقاربة وظيفية للغة ترتكز على استعمالها، وبالتالي تلتحق بالتيار الوظيفي[1].

1.2.1. الوظيفية المعرفية

بالنسبة للوظيفيين، من الضروري دراسة اللغة في شروطها الطبيعية، دون تجاوز الطبيعة الحقيقية للمكنزمات التي تهدف إلى تمثيل المعارف وإيصال هذه المعارف الممثلة، مما يفترض الانطلاق من الملفوظات الحقيقية المنتجة في سياق الاتصال، وليس من الملفوظات المولدة مجردة عن السياق. هذه المنهجية نابعة من تصور للنحو بأنه نمط خاص لتشفير اللغة مرتبط باستعمالها، ويسمح بالمعالجة في أفضل الحالات. النحو يشكل نظاما لمعالجة المعلومة متكيف بوجه خاص مع قيود الاتصال اللساني، ويمثل –بوجه محتمل– آخر ما توصلت إليه سيرورة التطور لبناء نظام اتصال إنساني[2].

2.2.1. الأنحاء المعرفية

تقوم الأنحاء المعرفية على المبادئ المنهجية السابقة نفسها، لكنها تمنح دورا معرفيا بنائيا للاستعمال في إطار نموذج حركي لاشتغالية اللغة. حسب هذا التصور، المفاهيم: الكفاية/الأداء، النظام اللساني (التجريد)/ الملفوظات الخاصة (التجربة) لا يمكن الفصل بينها، والدليل على ذلك أمران، الأول: أن المتكلم ينشئ نظامه اللساني انطلاقا من التجربة المكررة للإنتاج والفهم، من خلال سيرورة التجريد؛ إذ النشاط اللغوي هو نشاط خلاق وحركي من أجل بناء مفهمات، يتم في هذا النشاط إعادة بناء قيمة الوحدات اللسانية وتسويتها دوريا.

---

[1] ينظر: Fayol (Michel), Production du langage, op.cit, p. 72

[2] ينظر: *Ibid*, p. 73

الأمر الثاني: أن المكونات اللسانية المعجمية والنحوية لا تشكل بنيات صورية أو مضامين مرجعية بسيطة، ولكنها هي نفسها مفهمات؛ فهي تمثل في الوقت نفسه سيرورات الترميز أي ميكنزمات بناء التمثيلات، ونتيجة هذه السيرورات التي هي نفس التمثيلات في آن معا. وبالتالي، لا يمكن التمييز بين مستوى بنوي ومستوى للمعالجة: كل المكونات اللسانية –بما فيها النحوي– تشكل عمليات لمعالجة التمثيلات. هذه العمليات التي تكيفها التجربة تكشف عن مكنزمات عامة تتدخل في أنشطة معرفية أخرى[1].

نظرا لاعتمادها على المفهمات أكثر من اعتمادها على المكونات اللسانية، لا توفر هذه النظريات النحوية المعرفية مباشرة نماذج لإنتاج اللغة. نماذج **تالمي Talmy** و**لانقاكر Langacker** تسمح بالمقابل بوصل التمثيلات بالمكونات التي تبنيها. لقد بين **تالمي Talmy** خاصة بأن اللغة يحكمها ميكنزم تنظيم تراتبي يشبه نظام الصورة الفوتوغرافية التي تحتوي على الصورة المستهدفة وخلفية الصورة؛ ويشرح ذلك بأن بناء التمثيلات اللسانية ينتظم باستمرار على المبدأ التراتبي للصورة الفوتوغرافية حيث يؤخذ مفهوم ما كمرجع ثابت (خلفية الصورة) من أجل إدخال مفهوم آخر بحاجة إلى التحديد والذي يشكل موضوع هذا البناء (الصورة المستهدفة)[2].

تتحدد وظائف الصورة وخلفية الصورة بواسطة خصائص معرفية للعناصر المستخدمة (حضور العنصر في السياق، الحجم والدوام النسبيان، درجة التعقيد الهندسي...)، وتتحمل تلك الوظائفَ مكونات لسانية مختلفة المستويات. يمكن أن يفسر هذا اختلاف درجة مقبولية

---

[1] ينظر: Fayol (Michel), Production du langage, *op.cit*, p. 74

[2] ينظر: Fayol (Michel), Production du langage, *op.cit*, p. 75

الملفوظات رغم أنها سليمة التركيب والدلالة. فالموارد الرمزية في اللغة توفر عددا كبيرا من الصور البديلة في وصف مشهد معين، وبالتالي فتعدد الجمل أو أشكال التعبير عائد إلى تعدد زوايا التصوير المنصبة على الواقعة الواحدة دون أن تكون الجملة الواحد مشتقة من جملة أخرى، وعلى هذا لا وجود حسب **لانقاكر Langacker** للبنية العميقة ولا لتحويلات. مثال ذلك الجملتان الآتيتان:

أهدى زيد كتابا إلى عمرو.

- أهدى زيد عمرا كتابا.

إن هاتين الجملتين حسبه مختلفتان معنى من حيث تتحقق في الواحدة منهما صورة تختلف عن تلك المتحققة في قرينتها. فهما صورتان مختلفتان لواقعة واحدة هي الإهداء، ومضمون الواقعة هو نفسه فيهما (زيد، عمرو، كتاب)، والاختلاف بينهما كائن في ما يكون في الواحدة منهما من بروز نسبي لواحد من مظاهر الموقف المركب الذي تمثله الواقعة[1].

لقد عمم هذا الباحث فكرة الانتظام التراتبي للتمثلات اللسانية على شكل "صورة مستهدفة" و"خلفية الصورة"، وافترض بأن بناء الدلالة يمر عبر تخصيص مختلف أبعاد المعنى التي تملك ميكنزمات معرفية عامة. أحد الميكنزمات اللسانية هو **العرض *profiling*** يتمثل في وضع المعروض (الصورة) على أساس (خلفية الصورة). فالمعروض محط العناية ومركزها في إطار الأساس، وهو كذلك ما تحيل عليه العبارة من الأبنية الفرعية المكونة للأساس. مثلا: مفردة "خال" (أخو الأم) لا يمكن أن نتصورها ما لم يرتكز على أساس أو إطار الروابط

[1] ينظر: الزناد (الأزهر)، نظريات لسانية عرفنية، مرجع سابق، ص. 100.

الأسرية، فهو يستحضر شبكة كاملة استحضارا طبيعيا ولا يفيد معناه أي ذاك الشخص بعينه ما لم ينشدَّ إلى إطار يشمله وسائر العناصر المكونة لشبكة الأسرة[1].

الأنحاء المعرفية توفر إذن أسسا جديدة من أجل إنشاء نموذج لمعالجة اللغة، يقوم على ميكنزمات حركية لبناء التمثيلات التي يمكن -ولو جزئيا- أن ترتبط بمكنزمات معرفية أكثر تعميما. نموذج الإنتاج النابع من هذا التيار يحتاج إلى المزيد حتى يتم إنشاؤه، لكن معالمه بدأت تظهر مختلفة تماما عن تلك الخاصة بالنحو التوليدي[2].

2. النماذج اللسانية النفسية

يعد إنتاج الكلام في التصور اللساني النفسي حركة معاكسة لفهم الكلام، حيث تتم هذه العملية بتحويل أفكار المتكلم إلى عبارات. لقد أثارت هذه العملية تساؤلات عدة عند الباحثين: كيف يتسنى للمتكلم اختيار طريقة معينة من بين طرق كثيرة متاحة للتعبير عن تلك الأفكار ؟ ثم بعد اختيار بنية معينة، كيف يتم تنظيم الكلمات، وكيف يتم حساب التعيينات الصرفية والتركيبية من نوع وعدد وغيرهما ؟

يكاد علماء اللسانيات النفسية يجمعون على وجود ثلاث مراحل أساسية تتم فيها عملية إنتاج الكلام، وهي:

- مرحلة مفهمة الرسالة التي يراد التعبير عنها: وتتحدد فيها المفاهيم والأفكار المطلوب التعبير عنها لفظا حيث يتعلق الأمر بلغة الفكر التي هي مستقلة عن اللغة التي يعبر بها المتكلم.

---

[1] ينظر: نفسه، ص. 106-707.

[2] ينظر: Fayol (Michel), Production du langage, *op.cit*, p. 78

- مرحلة معجمة الرسالة: حيث يجري اختيار الكلمات الملائمة للمفاهيم وللأفكار التي يرغب المتكلم في التعبير عنها من أجل بلوغ بنى دلالية وتركيبية وفونولوجية وصرفية متماسكة. تسمى هذه المرحلة بالدخول المعجمي الذي يتم في مرحلتين فرعيتين:
  - مرحلة التشفير الدلالي والتركيبي: تتمثل في استرجاع اللمات، وتليها مرحلة التشفير الصرفي الفونولوجي التي تتمثل في استرجاع اللوسيمات[1].
- مرحلة مفصلة الرسالة: وهي المرحلة الأخيرة التي تقود إلى النطق بها.

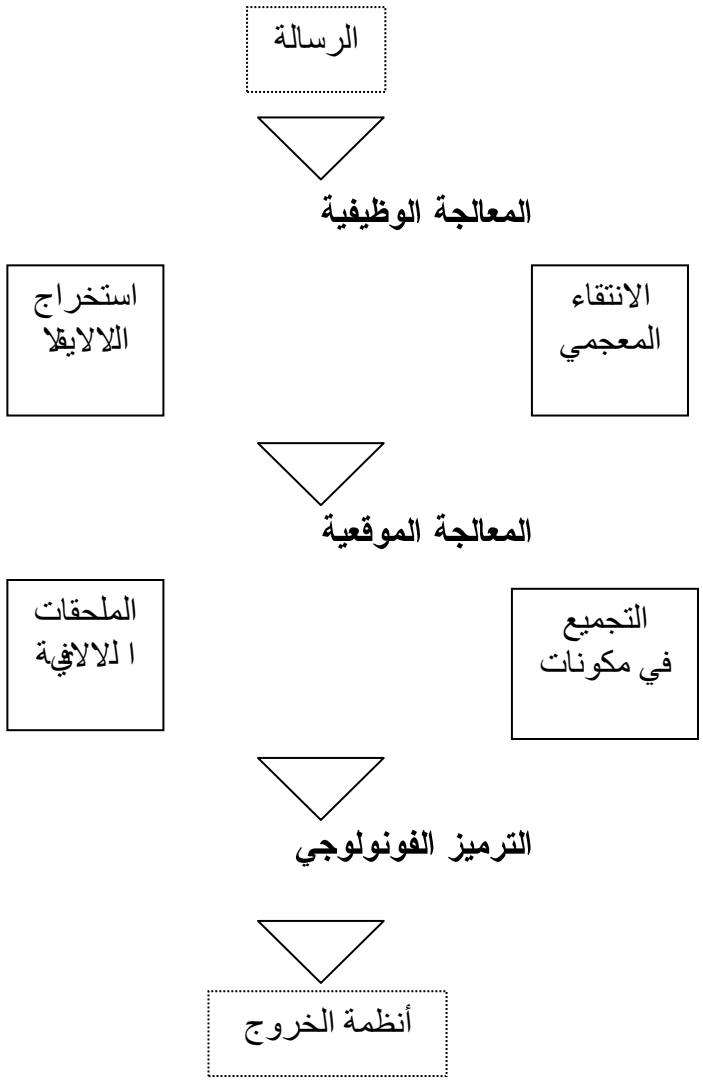
1.2. معمارية إنتاج الكلام

من أهم النماذج المقترحة والأكثر تطورا في مجال إنتاج الكلام نموذجا **لفيلت Levelt وديل Dell**.

1.1.2. مسلمات نموذج لفيلت Levelt

يمكن تمثيل هذا النموذج في الشكل الآتي:

[1] اللمات هي كيانات مجردة تتمثل في الخصائص الدلالية والتركيبية لكلمة ما. أما اللوسيمات فهي كيانات مجردة تتمثل في الخصائص الفونولوجية لكلمة ما (عدد المقاطع، وطبيعة الفونيمات). ينظر: Ferrand (Ludovic), « La production du langage eee '' eeeeeeee », *in* Psychologie française, mars 2001, tome 46 – n°1, PUG, p. 5.



الشكل9: نموذج إنتاج الكلام المنطوق حسب لفيلت Levelt[1]

حسب هذا النموذج، تتم عملية إنتاج الكلام عبر أربع مراحل من المعالجة:

- المرحلة الأولى: معالجة الرسالة. حيث يقوم المتكلم في هذه المرحلة بصياغة مضمون الرسالة التي يريد إبلاغها، والمعنى الذي يريد أن يوصله للسامع.
- المرحلة الثانية: المعالجة الوظيفية. يقوم المتكلم خلال هذه المرحلة باختيار المعجم المناسب، وكذا المعاني التركيبية (فاعل، مفعول...) التي يريد إضفاءها على الكلمات المختارة من المعجم.

[1] ينظر: Lemaire (Patrick), Psychologie cognitive, Deboeck, Paris, 1999, p.337

- المرحلة الثالثة: المعالجة الموقعية. في هذه المرحلة، يقوم المتكلم بوضع الكلمات المختارة ضمن ترتيب محدد من أجل تكوين الجملة، وكذا يمنح تلك الكلمات خصائصها وملحقاتها التصريفية.
- المرحلة الرابعة: المعالجة الفونولوجية. وأخيرا، يعمد المتكلم إلى اختيار الأصوات التي تتشكل منها الكلمات داخل الجملة المؤلفة، وكذلك يعمل مع سائر البرامترات المتصلة بالكلام المنطوق كالتنغيم والإيقاع...

يتميز هذا النموذج بكونه نموذجا معالجا للمعلومة بشكل تسلسلي، مما يعني أنه يتخذ عدة سيرورات يمكن أن تفحص في المخبر، ومن خصائصه أيضا اهتمامه بالمعلومات الدلالية والتركيبية بالإضافة إلى المعلومات الفونولوجية[1].

2.1.2. مسلمات نموذج ديل

يقترح **ديل Dell** من جهته نموذجا ترابطيا لإنتاج الكلام، وتتم هذه العملية أيضا في أربع مراحل:

- المرحلة الأولى: المعالجة الدلالية. في هذه المرحلة يقوم المتكلم بصياغة معنى الرسالة التي يريد إيصالها.
- المرحلة الثاني: المعالجة التركيبية. يضع المتكلم في هذه المرحلة مخططا للبنية النحوية التي سيضمنها قوله؛ أي ينظم الكلمات لتشكيل الجمل.

[1] ينظر: Lemaire (Patrick), Psychologie cognitive, *op.cit*, p. 338

- المرحلة الثالثة: المعالجة الصرفية. يقوم المتكلم خلالها بتصميم المورفيمات؛ أي يعمل على اختيار العناصر التي تعيّن حدثا أو أكثر، وإن كان ذلك الحدث يتم في زمن القول أو في زمن آخر...
- المرحلة الرابعة: المعالجة الفونولوجية. يختار المتكلم في هذه المرحلة الأصوات التي يستعملها، وكذلك التنغيم والإيقاع...[1]

3.1.2. الاختلاف بين النموذجين

يظهر أن النموذجين متفقان في عدة نقاط إلا أن بينهما اختلافات جذرية. فحسب **ديل**، يعمد المتكلم إلى صياغة تمثيل عند كل مرحلة من مراحل المعالجة، ويعتبر تلك التمثيلات ضرورية في كل لحظة من لحظات المعالجة. كما يرى ديل أن عمليات المعالجة تتم بطريقة متوازية على عكس عمليات المعالجة في نموذج **لفيلت** التي تتم بطريقة متسلسلة؛ إذ لا تنطلق فيه عملية المعالجة الموالية إلا بعد انقضاء سابقتها، أما في منوال **ديل** فتتداخل عمليات المعالجة وتتفاعل فيما بينها ويؤثر بعضها في بعض. هناك فرق جوهري آخر بين النموذجين يكمن في أن منوال **ديل** يتضمن ميكنزمات عامة للمعرفة الإنسانية. إن الميكنزم الأساسي الذي يتدخل في عملية الإنتاج النطقي هو ميكنزم الإنشاط[2]. وعليه، يدفع هذا الميكنزم إلى استحضار المعلومة واستذكارها، وينشط فيها كل مفهوم والمفاهيم المرتبطة به وينتشر فيها[3].

---

[1] ينظر: Lemaire (Patrick), Psychologie cognitive, *op.cit*, p. 339

[2] وهو ميكنزم يفترض الباحثون فيه بأن التمثيلات المعرفية تتحدد بحالتين: حالة راحة (حالة اللانشاط)، وحالة نشاط وهي حالة تقبل تبدلات تظهر في مستويات النشاط المختلفة، وهذا الانتقال توفره سيرورة التنشيط. ينظر: Gineste (M-D) ; Le Ny (J-F), Psychologie cognitive du langage, DUNOD, Paris, 2002, p.149

[3] ينظر: Lemaire (Patrick), Psychologie cognitive, *op.cit*, p. 489

2.2. مراحل إنتاج الكلام

انطلق **جاريت Garrett** في دراسته لإنتاج الكلام من أربع مستويات للمعالجة، وقد توصل إلى هذه النتيجة انطلاقا من تحليله للأخطاء العفوية، وهي: معالجة الرسالة، المعالجة الوظيفية، المعالجة الموقعية، المعالجة الصوتية. تسمى المعالجتان الوظيفية والموقعية بمعالجتي التشفير النحوي، ويصنفها الباحثون في مجموعتين: الأولى مهمتها إنشاء تمثيل للمستوى الوظيفي حيث يتم توزيع الوظائف التركيبية (الفاعل، المفعول...) على الوحدات المعجمية المخزنة في الذهن، والثانية مهمتها إنشاء تمثيل للمستوى الموقعي والذي يتضمن ترتيب الكلمات بإدماج أشكالها داخل بنية تركيبية.

لقد خرج **جاريت** بذلك التمييز بين المستويين التمثيليين الوظيفي والموقعي من خلال مقارنته بين ثلاثة أصناف من الأخطاء يقع فيها تبادل عنصرين، في المثالين:

That wasn't a TIGER, that was a LION.

لم يكن ذلك نمرا، بل كان أسدا

On a Slot Holdering iron. (hot/ soldering)

على (ســـ)ديد (حـــ)ـــاخن للتلحيم

العناصر المتبادلة مكتوبة بأحرف كبيرة؛ في المثال الأول يتم تبادل العنصرين « tiger » و« lion »، في حين يحدث تبادل بين صوامت في المثال الثاني: /h/ من « hot » و/sl/ من « soldering ».

التعليق على هذا النوع من الأخطاء التي قد يقع فيها المتكلمون باللغة الإنجليزية تلقائيا يبين أن تبادل عنصرين غير ممكن إلا داخل وحدة المعالجة نفسها، كما أنه توجد مستويات

مختلفة للمعالجة تتضمن وحدات مغايرة. فالاختلافات النظامية التي تخص القيود البنوية على موقع الأخطاء هي التي دفعت **جاريت** إلى التفريق بين المستويين الوظيفي والموقعي. مقارنة مع التبادلات الصوتية التي تقع في المثال الثاني، تنتمي الكلمات المتبادلة في المثال الأول عادة إلى مركبات (أركان) مختلفة، مما أدى بهذا الباحث إلى الافتراض بأن وحدة المعالجة هي الجملة على المستوى الوظيفي، والكلمة على المستوى الموقعي[1].

بالإضافة إلى ذلك، تخص مبادلة الكلمات عادة عناصر من المقولة النحوية نفسها أكثر من المبادلة التي تحدث بين الأصوات. المقولة النحوية[2] تتدخل إذن عند إدماج الكلمات في الجمل.

هناك صنف آخر من الأخطاء يمكن ملاحظتها في المثال الثالث:

YOU're too good for THAT

عوض: THAT's too good for YOU

"أنت مناسب تماما لذلك" عوض: "ذلك مناسب لك تماما"

الكلمتان « you » و« that » متبادلتان، لكن مطابقة الفعل للاسم الذي يحتل الموقع تبدو صحيحة على الرغم من أنه ليس الاسم المنتظر أن يحل في هذا الموقع.

إذن يُحتمل تبادل الكلمات في المرحلة التي ستُحدّد فيها العلاقات التركيبية بين العناصر المعجمية، ويُتوقع تبادل الأصوات في المرحلة التي ستُخصص فيها الأشكال الفنولوجية

---

[1] ينظر: Fayol (Michel), Production du langage, *op.cit*, p. 108

[2] **المقولات النحوية *catégories grammaticales*** هي أصناف العناصر التي تظهر في المحيط التركيبي نفسه، وترتبط فيما بينها بعلاقات خاصة. ينظر: Dubois (Jean) et *al*., Dictionnaire de linguistique et des sciences du langage, *op.cit*, p. 78.

للعناصر ويُحضّر تنظيمُها التسلسلي. في هذه المرحلة، تنشأ المكونات المركبية في شكل سلمي أو هرمي حيث تحدد العلاقات الخطية والتراتبية بين المكونات. هذه الفرضية يثبتها التحليل الصوري للوقفات في الخطاب والأخطاء المرتكبة.

تحليل المبادلات بين الجذور أيضا في المثال الرابع يشير إلى الملحقات الصرفية بالعناصر المعجمية في المرحلة الموقعية.

المثال الرابع:

I thought the PARK was TRUCKed

عوض: I thought the TRUCK was PARKed

"أظن أن السيارة قد رُكنت"

الجذران « park » و« truck »، متبادلان، لكن العلامة الصرفية في الفعل park لم تشارك في عملية المبادلة، لذا تبدو متصلة بالاسم « truck » في الجملة المنتجة بالفعل.

مبادلة الجذور من جهة تخضع لنفس القيود الخاصة بأخطاء الأصوات (العناصر تنتمي إلى نفس الركن، وهي مقولات تركيبية مختلفة). ومن جهة أخرى، اللاحقة ***ed***- موقعها صحيح لكنها مرتبطة بعنصر غير مناسب. هذه الخصائص تشير إلى أن الملحقات الصرفية توضع بمعزل عن جذور الكلمات[1].

ولعل ذلك قد يكون أظهر في اللغة العربية حيث تكون الملحقات الصرفية أحيانا عبارة عن وزن أو صيغة صرفية مثلا يقابل تلك الملحقة الصرفية في الإنجليزية الوزن فَعَلَ في العربية:

---

[1] ينظر: Fayol (Michel), Production du langage, *op.cit*, p. 108

الفعل	الماضي
lived	-ed
سَكَنَ	فَعَل

ويبدو أن صياغة الفعل الماضي على وزن 'فَعَل' سابقة على إحلال العنصر المعجمي المتمثل في مادة (ف ع ل) فيه.

يقترح **جاريت** أن تنتمي عناصر الصف المغلق التي تتمثل في الكلمات الوظائف ومختلف العلامات داخليا إلى الإطار النحوي، وأن عناصر الصف المفتوح التي تتضمن الأسماء والأفعال والصفات وغيرها ينبغي أن توصل بالإطار النحوي بعد أن تزوَّد بوظيفة نحوية. وبيّن كذلك أن الأخطاء المتعلقة بالكلمات الوظائف كالمحددات وحروف الجر هي أقل ترددا من تلك التي تتعلق بسائر الكلمات في الصف المفتوح.

انطلاقا من تحليله لأخطاء الخطاب، أقام **جاريت** نموذجا لتنظيم مستويات المعالجة في الإنتاج العادي للكلام. معالجات التشفير النحوي تستقبل في المدخل التمثيل المفهومي للمعنى الذي يريد أن يعبر عنه المتكلم (الرسالة)، وهي تنقسم إلى مستويين وظيفي وموقعي. في المستوى الوظيفي، تمنح الوظائف التركيبية للعناصر المعجمية التي تترجم مقاصد المتكلم؛ وفي المستوى الموقعي، توضع بنية موقعية (إطار تركيبي) يحتوي على مواقع شاغرة أي يحتوي على متغيرات يمكن أن تملأها مواضيع خاصة[1].

[1] ينظر: Fayol (Michel), Production du langage, *op.cit*, p. 109

هذه المواقع الشاغرة ستحتلها كلمات وعلامات للزمان والشخص والعدد للأفعال، والعدد والجنس للأسماء. وبالنسبة للعربية، تحتل تلك المواقع كلمات وعلامات قد تكون لاحقة بالفعل مثل تاء التأنيث الساكنة الملحقة بالماضي (كتبَتْ)، أو سابقة عليه مثل الحرف 'قد': قَدْ كَتَب. ويضاف إلى ذلك في اللغة العربية أن أشكال الفعل أو صورته (الميزان الصرفي) لها دلالة على الزمان والشخص الذي يقوم بالفعل، أما العدد فليس علامة تلحق الفعل في العربية. وبالنسبة لعلامات الأسماء يمكن أن نضيف الإعراب والتعريف والتنكير بواسطة علامات تلحقها تكون في شكل حركات (صوائت) أو حروفا (نصف صوائت أو شبيهة بالصوائت).

1.2.2. المعالجات المعجمية

بكلا المستويين الوظيفي والموقعي ترتبط مرحلة للدخول المعجمي، وبالتالي هناك مرحلتان للدخول المعجمي: 1- استرجاع اللمات وهي العناصر المعجمية المجردة والمستقلة عن صورتها الفونولوجية التي تترجم المفاهيم المعجمية، 2- استرجاع اللوسيمات وهي الصور الفونولوجية للمفردات. تتدخل الأولى عند إنشاء التمثيل الوظيفي، فتمنح الوظائف التركيبية (فاعل، مفعول...) للمات. وتتدخل الثانية عند إنشاء التمثيل الموقعي حينما تملأ اللوسيمات المواقع الشاغرة للإطار الموقعي (التركيبي)[1].

عند إنتاج الجمل، كل اللمات التي تنتمي إلى وحدة المعالجة نفسها (الجملة) تنشط بالموازاة مع المرحلة الوظيفية. في المرحلة الموقعية، تدخل اللوسيمات في البنية التركيبية ركنا

[1] ينظر: Ferrand (Ludovic), « La production du langage rrrrrrrrr ddeerrrrrrr », *op.cit*, p. 7

ركنا. في المستوى الوظيفي، تمنح الوظائف التركيبية للمات[1]. سأتطرق في الفقرة الموالية إلى فحص العوامل التي تؤثر في تعيين الوظائف التركيبية.

2.2.2. المستوى الوظيفي

في هذه المرحلة، تمنح الوظائف التركيبية للعناصر الدلالية (اللمات الأكثر إنشاطا)، عندما تمنح تلك الوظائف تكون المرحلة التالية محددة سلفا. إذا كانت الجمل تصف الواقعة نفسها، وكانت بنياتها التركيبية مختلفة ترتب عن ذلك أن تُمنح للعناصر وظائف تركيبية مختلفة. وعليه، فإن الصورة الفونولوجية للكلمات (اللوسيمات) لا يمكن أن تؤثر في الصورة التركيبية للجمل إلا عندما يُراجع التمثيل الوظيفي.

3.2.2. المستوى الموقعي

حينما يكتمل إنشاء التمثيل الوظيفي، يتم إرساله إلى المعالج التركيبي الذي يقوم بإنشاء بنية مكونات اعتمادا على المعلومات المتوفرة. في هذا المستوى أيضا، تُدخل العلامات (الملحقات التصريفية).

- **تجميع المكونات**

في المستوى الوظيفي، تعيين الوظيفة التركيبية متأثر ببروز عنصر من عناصر الملفوظ، والذي يُمنح بالنتيجة الوظيفة التركيبية الأولى في الترتيب (الفاعل ثم المفعول...). عندما تُمنح الوظائف التركيبية تكون البنية المكونية كاملة التحديد[2].

---

[1] ينظر: Fayol (Michel), Production du langage, *op.cit*, p. 113

[2] ينظر: Fayol (Michel), Production du langage, *op.cit*, p. 116

في المرحلة الموقعية، معالجات إنشاء بنية المكونات توظف البنية المقولية الفرعية للأفعال. التأشير بالعلامات ليس مستقلا عن إنشاء هذه البنية المكونية، وإن تضمن إيجاد تفاصيل لأدنى مستوى من البنية.

- **التأشير بالعلامات**

في أغلب اللغات، تجب المطابقة بين الفعل وفاعله. في اللغة العربية، ليس في الفعل علامة تدل على العدد إلا على لغة "أكلوني البراغيث" -كما يسميها النحاة- اللهم إلا إذا تقدم الفاعل على الفعل، وفي هذه الحال يعد مبتدأ، أما الفاعل فهو الضمير العائد عليه والذي يأخذ أشكالا عدة بحسب العدد، فلا يشبه ذلك المطابقة بين الفعل والفاعل في اللغة الفرنسية مثلا.

بكى الطفل ← الطفل بكى / l'enfant pleure

بكى الطفلان ← الطفلان بكيا / les enfants pleurent

بكى الأطفال ← الأطفال بكوا / les enfants pleurent

في اللغات الغنية على مستوى الملحقات التصريفية كالإيطالية، تكون المطابقة من حيث الجنس بين النعت ومنعوته أو بين الاسم والضمير العائد عليه ضرورية. وبالتالي، تتعلق معالجات المطابقة بين مختلف العناصر بالعلاقات التركيبية، وتنتمي بالنتيجة إلى التشفير النحوي[1].

[1] ينظر: Fayol (Michel), Production du langage, *op.cit*, p. 120

3.2. نمذجة إنتاج الكلام

إن النماذج الحاسوبية لإنتاج الكلام هي بناءات معقدة ومتنوعة؛ فهي تقدم المعلومات والسيرورات المرتكزة على فرضيات شتى. بالإضافة إلى ذلك، لا تتبنى تلك النماذج الأهداف نفسها؛ إذ لا تتفق فيما بينها على نوع البيانات (مثلا: محاكاة نماذج الأخطاء أو زمن الإجابة)، وعلى الطريقة (محاكاة بسيطة لنتائج عامة أو توزيعات كاملة أنتجها الفاعلون).

في هذا السياق، يسمح تحليل البرمجة الحاسوبية لتلك الفرضيات بتفكيك النماذج المقترحة وميكنزماتها، وبالتالي يمكن تحديث الأسباب التي تدفع إلى محاكاة واقعة أو عدمها. كما أن الفهم الجيد لدلالة الصياغة الرياضية يتيح التأكد من موافقة الفرضية المعرفية التي تعمل على تمثيلها. وفي الأخير، يشكل هذا التحليل مرحلة أساسية للمقارنة بين النماذج التي تمت صورنتها باستخدام صورانيات (أمثلات) مختلفة.

وفي هذا الصدد، تتنوع البرمجيات الحاسوبية المستخدمة؛ ففي علم النفس اللساني، تهتم دراسة إنتاج الكلام بالسيرورات التي تسمح باسترجاع المعلومة اللسانية من الذاكرة من أجل التعبير عن رسالة ما. أبرز قاسم مشترك في هذا الحقل الدراسي هو كون المعالجات اللسانية المدروسة أساسا تحت تحكم المعلومات الدلالية والمفهومية. هذا التحديد يقصي بعض الحالات التي يتم فيها إنتاج الكلام كالقراءة مثلا التي تكون فيها المعلومات اللسانية (التمثيلات الإملائية) هي المتحكمة في السيرورات المدروسة، وليست المعلومات الدلالية والمفهومية[1].

---

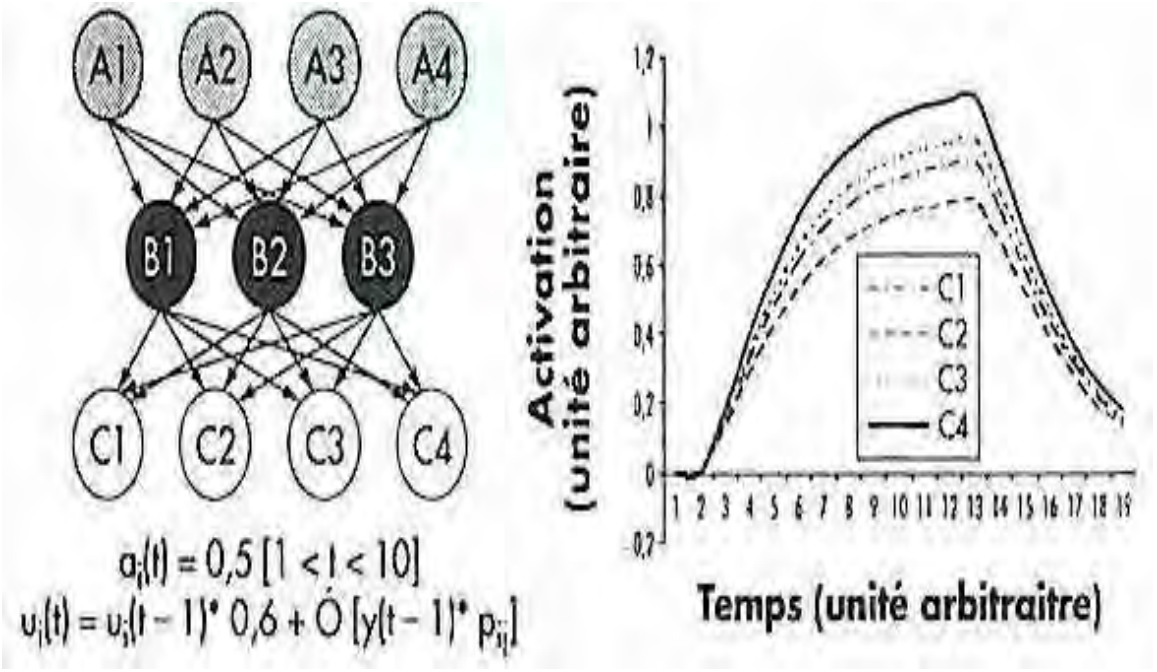
[1] ينظر: Alario (François-Xavier) et *al*., Psycholinguistique cognitive, De Boeck, Bruxelles, 1re éd., 2004, p. 145.

1.3.2. العناصر المكونة للنماذج الحاسوبية

في أغلب الحالات، ترتكز النماذج الحاسوبية لسيرورات إنتاج الكلمات (المفردات) على مجموعة وحدات للتمثيل هي العصبونات الاصطناعية. من وجهة نظر رياضية، هذه الوحدات هي متتاليات عددية $\mathbf{U_t}$ مؤشرها هو البعد الممثل للزمن (متقطع) حيث يتطور النموذج. هذه الوحدات تمثل المعلومات اللسانية المميزة من أجل تحقيق السلوك المدروس –المعلومات اللسانية مثلا–، ولكن يمكن أن تمثل معلومات ليست مؤولة نفسيا بصفة مباشرة.

ثم إن القيم التي تأخذها المتتاليات العددية في فترة زمنية والتي تتوافق مع وحدة ما تشكل درجة إنشاط المعلومة التي تمثلها تلك الوحدة. ترتبط مستويات الإنشاط تلك فيما بينها بواسطة نظام المعادلات التي تسيّر تطورها بدلالة الزمن. في الحالة العامة، قيمة إنشاط وحدة في زمن معين هي توليف قيم المتتاليات الأخرى في الزمن السابق، ويمكن أن تتدخل معايير أخرى. المعادلات التي تربط بين إنشاطات الوحدات تترجم الفرضيات المعرفية التي تُراد برمجتها.

إذن بالاعتماد على نظام المعادلات وانطلاقا من مجموعة قيم أولية، بالإمكان تعقب تطور الإنشاطات لكل المعلومات بدلالة الزمن (كما هو مبين في الشكل10).



الشكل10: نموذج بينة شبكة عصبونات

الشكل على اليمين هو تمثيل تخطيطي لنموذج مصغر للإنتاج تمت صياغته للتوضيح. هذا النموذج المصغر للإنتاج يتقاسم مع عدد من النماذج بنية من 03 مستويات، دون ترابط بيني للمستويات، ودون ترابط بين المستوى الأول والثالث (A→C). أما على اليسار، فيتضمن مثالا لمستوى إنشاط وحدات الخروج (معايير بث الإنشاط مثبتة عشوائيا)[1].

إن نماذج التوليد، التي اقترحها اللسانيون على اختلاف مشاربهم سواء أكانت نماذج للكفاية أو للأداء، والتي عمل الباحثون في اللسانيات النفسية على دراسة مختلف المعالجات التي تتضمنها تلك العملية، تمثل مرحلة أساسية لإنشاء برمجيات تساعد على إنجاز تلك العملية بطريقة آلية.

[1] ينظر: Alario (François-Xavier) et *al*., Psycholinguistique cognitive, *op.cit*, pp. 146-147

لكن التشغيل الآلي لعملية التوليد يقتضي، إضافة إلى ذلك، توفير مجموعة من الوسائل الرياضية والمنطقية المناسبة التي تعين على صورنة تلك النماذج وتسهيل حوسبتها ضمن مجال تطبيقي وفي إطار المعالجة الآلية للغات الطبيعية هو التوليد الآلي للنصوص أو الجمل.

الفصل الرابع:

التوليد الآلي للجمل

الفصل الرابع: التوليد الآلي للجمل

يعتبر التوليد الآلي للجمل أحد الميادين المهمة للمعالجة الآلية للغات الطبيعية، نظرا لارتباطه بتطبيقات عدة ولحاجة جل التطبيقات إليه كالترجمة الآلية، وقواعد البيانات. أحاول أن أتعرض في هذا الفصل إلى توصيفه وبيان المشاكل والصعوبات التي يتضمنها، وأن أذكر أبرز المقاربات والمناهج المستخدمة في هذا الميدان.

1. التوصيف

التوليد الآلي للجمل أو النصوص هو أحد المجالات التي تتناولها المعالجة الآلية للغات الطبيعية وفرع اللسانيات الحاسوبية، حيث يتم فيه توليد صورة أو شكل لساني انطلاقا من معطيات غير لسانية أي انطلاقا من تمثيل صوري لمحتوى دلالي. وبالتالي، يتضمن ميدان التوليد تطوير أنظمة قادرة على إنتاج نصوص أو ملفوظات متسقة، سليمة ومفهومة. الأنظمة الحاسوبية المنجزة من أجل توليد النصوص بطرق آلية ترجع إلى سنوات 1970.

يقع هذا الميدان في مفترق الطرق بين المعلوميات والذكاء الاصطناعي واللسانيات والعلوم المعرفية. في الواقع، تكمن قيمة نظام التوليد أو المولد في قدرته على معالجة ثلاثة أنواع مختلفة من المشاكل:

- يتعلق المشكل الأول بالتفاعل بين الإنسان والآلة: ما هي الوسيلة المثلى كي تنقل الآلة المعلومة إلى الإنسان ؟ كيف ينبغي لها أن تتصل بالإنسان، وكيف يمكن أن يبرمج هذا السلوك ؟ هذا النوع من الأسئلة يوجه إلى حقل المعلوميات.

- ويتصل المشكل الثاني بصورنة الاتصال: كيف باستطاعتنا صورنة القيود التداولية والدلالية والتركيبية والنفسية اللسانية ؟ ما هو دور السياق في عملية الاتصال ؟ هذا النوع من الأسئلة موجه إلى حقول اللسانيات والفلسفة والعلوم المعرفية.
- أما المشكل الأخير فمرتبط بالنمذجة ومحاكاة الحاسوب للذكاء الإنساني: كيف تتسنى ترجمة التمثيلات الرقمية الخاصة بالمعلومة إلى تمثيلات متلائمة مع الإنسان ؟ وبالتالي، من خلال إثارته لهذا المشكل، يتقاطع ميدان التوليد الآلي مع حقل الذكاء الاصطناعي[1].

إن مجال التوليد يقابل مجالا آخر يبدو متناظرا معه هو مجال "فهم اللغة الطبيعية"؛ فإذا كانت عملية التوليد تجري بتحويل المعنى إلى نص، فإن عملية الفهم تبدأ من النص لتستخرج المعنى الذي يتضمنه. لو أجرينا مقارنة بين العمليتين لوجدنا أن **دخل *input*** التوليد يتميز بتعدده واختلافه (المعنى المعجمي، الدلالي، التداولي، الصرفي، التركيبي...) وأن **خرجه *output*** هو واحد (النص نفسه). أما في عملية الفهم فنلاحظ العكس إذ يكون الدخل (النص) واحدا والخرج (المعنى) متعددا ومختلفا.

لذلك، فإن هناك العديد من المشاكل التي يمكن أن تواجه الباحثين في مجال الفهم قد تختلف نوعا ما عن المشاكل التي تصادفهم في مجال التوليد. ففي الحالة الأولى يعترضهم مشكل الغموض الذي يمكن أن يكتنف الجمل، والذي يتجاوزه المخاطَب المستقبل لتلك الجمل في الحالة الطبيعية والعادية باعتماده على قرائن سياقية (موقف التخاطب، وضعية المتكلم، الإطار العام للحوار...)، لكن الآلة تعجز عن إدراك تلك القرائن نظرا لتعددها وتعقدها. في

---

[1] ينظر: El Kassas (Dina), *Une étude contrastive de l'arabe et du français dans une perspective de génération multilingue*, thèse de Doctorat, Université Paris VII, 2005, p. 5.

حين أن المشكل الذي يلاقونه في عملية التوليد هو مشكل الاختيار أي ينحصر أساسا في كيفية تحديد نص ما من بين النصوص الممكنة حتى تكون خرجا للمعنى المطروح لدى الدخل، هذا المشكل يتجاوزه المتكلم عادة بتحديد أغراضه ومقاصده التي يريد إيصالها وتبليغها إلى المخاطب.

لم يكن الباحثون دوما يقيمون التقابل بين التوليد والفهم، فلطالما قابلوا مجال توليد النص بمجال تحليل النص فيما يخص الاتصال بين الإنسان والآلة، لكن **بولغير Polgaire** يعتقد أن هناك تفاوتا بين المصطلحين وعدم مناسبة؛ فالعملية المعاكسة للتحليل هي التركيب وليست التوليد مهما كان التطبيق الحاسوبي المستهدف، وذلك لأن توليد النص يفترض نشاطا أكثر تعقيدا من تحليل النص: توليد النص هو أن تنتج نصا انطلاقا من لا شيء أو بتعبير أصح انطلاقا من شيء لا يتمتع بصفة كونه نصا، إشارة إلى ما يتمتع به معنى النص من جذور نفسية وفكرية. وبالتالي فإن العملية المعاكسة لتوليد النص هي فهم النص أو تأويله حيث يكون المنطلق من النص وليس من تمثيله أو تمثيل معناه. يكوّن هذا إما مجموعة من المعارف الجديدة أو تحديثا للتمثيل الخاص بحالة العالم[1]. الجدول الموالي يظهر الفروقات الجوهرية بين هذه المصطلحات:

[1] ينظر: Polguère (Alain), *Structuration et mise en jeu procédurale d'un modèle linguistique déclaratif dans un cadre de génération de texte*, thèse de doctorat, Université de Montréal, 1990, pp. 6-7

	دخل	خرج
توليد	* معارف حول العالم * قواعد بيانات	نص
فهم أو تأويل	نص	* معارف جديدة حول العالم * قواعد بيانات جديدة
تركيب	تمثيل مضمون نص	نص
تحليل	نص	تمثيل مضمون نص

الجدول6: توليد/فهم في مقابل تركيب/تحليل

هذا الجدول يلخص الوصف الاصطلاحي لمفاهيم التوليد، الفهم، التركيب، التحليل حسب السمات التي تبدو مميزة، والتي ترتكز على خصوصيات المداخل والمخارج للنسق الذي تجري فيه هذه العمليات. لا توجد إشارة في المداخل والمخارج ضمن هذا الجدول إلى محددات للسياقات التي يمكن أن ترد فيها تلك العمليات من أجل أن تشمل تلك المداخل والمخارج كل السياقات الممكنة، كما أن النصوص تشمل كل منتج كلامي مكتوب أو منطوق، أما تمثيلاتها فتخص الجوانب التركيبية والدلالية والمفهومية وغيرها.

في حالتي توليد وفهم النص، تؤخذ في الاعتبار تلك العمليات التي تتم باستعمال عاديٍّ عند المتكلمين أو بصفة آلية عن طريق الأجهزة التي تنتج هذا الاستعمال. وبالتالي فإن دراسة تشغيل هذه العمليات لا ينتمي إلى النظرية اللسانية بل إلى ما يسمى بالعلوم المعرفية أو الذكاء الاصطناعي حيث يطبَّق برنامج يرتكز على مفاهيم مجردة، فيعطي تشغيله إعادة إنتاج العقل

الإنساني[1]. أما اللسانيات النظرية والوصفية فتقترح بالمقابل منوالات أو نماذج اللسان التي يمكن أن تستعمل عندما نريد محاكاة أو تقييس النشاط اللغوي؛ أي تتدخل تلك المنوالات في بنينة أو صياغة عمليات التركيب والتحليل كما تبيَّن سلفا.

إن التوليد الآلي للنصوص هو من بين الميادين الأكثر أهمية لأنه أساسي وجوهري بالنسبة للتطبيقات الحاسوبية مثل:

- الترجمة الآلية.
- قواعد البيانات.
- الأنظمة الخبيرة[2].
- الإنتاج الآلي للتقارير.

كما أن مجال التوليد الآلي للنصوص يشكل أهمية بالغة بالنسبة إلى اللساني؛ فتوليد النصوص يسمح له بأن يعالج إحدى المشاكل التي تبرز في اللغة وهي مشكل اختيار جملة من بين مجموعة من الجمل المترادفة أو المتشابهة من حيث مضمونها نحو:

- لا أستطيع الخروج في هذا الجو البارد.
- برودة الطقس تمنعني من مغادرة البيت.
- سألزم منزلي حتى ينتشر الدفء في الخارج.
- ...

---

[1] ينظر: Danlos (Laurence), Génération automatique de textes en langues naturelles, Masson, Paris, 1985, p. 47.

[2] النظام الخبير هو برنامج حاسوبي يبنى باستخدام أساليب الذكاء الاصطناعي بغرض محاكاة مهارات الخبراء في المشكلات المختلفة أو حلها. ينظر: معجم الحاسبات، مرجع سابق، ص. 74 .

إن مجال توليد النصوص يفضي إلى حل هذا المشكل اعتمادا على معارف لغوية محضة دون الحاجة إلى أن يرجع إلى معارف غير لغوية كما هي الحال في مجال الفهم والتحليل لأن المعنى المجرد لتلك الجمل المترادفة هو واحد لا يتعدد وإنما تتعدد الأساليب والصيغ التي تكتسي بها المعنى، لذلك يُطرح مشكل كيفية اختيار المتكلم لجملة محددة دون غيرها من الجمل. وعليه، ركزت **دانلوص Danlos** في نموذجها الخاص بتوليد النصوص على المعارف اللغوية، وبرهنت على متانة الصلة الموجودة بين اختيار المعجم والأبنية واستحالة تناولهما بطريقة منفصلة[1].

إن مشكل الاختيار يعترض الباحثين في كل مستوى من مستويات التمثيل:

- اختيار المحتوى: ما الذي نقوله ؟
- الاختيار المعجمي والتركيبي: كيف نقوله ؟
- الاختيار الخطابي: متى نقول قولا ما ؟
- العرض النصي أو اللفظي: القول في حد ذاته[2].

إن علاج مشكل الاختيار يتطلب الإحاطة بمجموعة من المعارف المسبقة أثناء عملية التوليد:

- معارف تتعلق بالحقل الذي تغطيه عملية التوليد (المهن، الصناعات، الأغذية...).
- معارف خاصة باللغة المستخدمة في ذلك الحقل.

---

[1] ينظر: Vuillaume (Marcel), Variation sur la référence verbale, google book, pp. 277-278

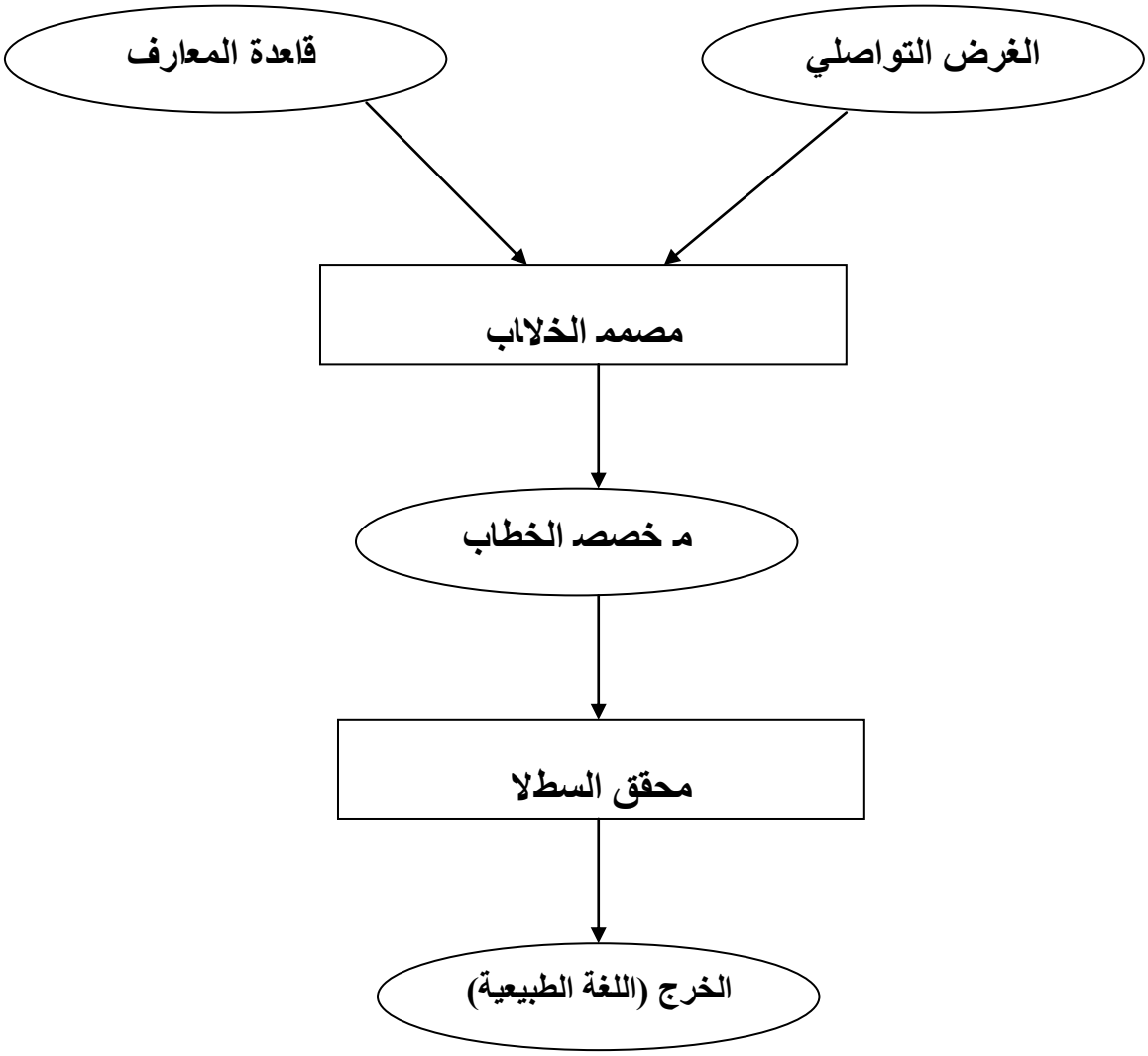
[2] ينظر: Mitkov (R.), The Oxford Handbook of computational linguistics, Oxford University Press, New York, 1st published, 2003, p. 286.

- معارف خطابية استراتيجية (محددات الأغراض والمقاصد أثناء التواصل).
- معارف تتعلق بالهندسة الحاسوبية لأن الهدف هو إنشاء برنامج حاسوبي يقوم بتوليد النصوص.
- معارف خاصة بعادات وضوابط مستخدم ذلك البرنامج الحاسوبي الذي يقوم بتوليد النصوص.

وعليه، يجب على البرمجة أن تراعي عدة عوامل منها القواعدية وعدم اللبس والانسجام، والغرض التواصلي المناسب، وأيضا الضوابط الاجتماعية والتخاطبية والتداولية. كما يشترط أن تكون المعلومات، التي تتضمنها التمثيلات الدلالية والتي توضع دخلا لنظام التوليد، ملاحظة لدى النص المنتج في الخرج. بالإضافة إلى ذلك، يكون المعجم ثريا والأبنية متنوعة والأساليب مقبولة عند المتكلمين[1]، لكن ذلك أمر في غاية الصعوبة كما يبدو؛ إذ يطلب من البرنامج أن يكون عقلا بديلا للعقل الإنساني، وليس برنامجا يحاول أن يلامس النشاطات اللغوية عند الإنسان من خلال إجراء عمليات مشابهة له.

يمثل **فندر ليندن Vander Linden** لمعمارية نظام التوليد الآلي للنصوص بالمخطط المبسط التالي:

[1] ينظر: Vuillaume (Marcel), Variation sur la référence verbale, *op.cit*, pp 277-278



الشكل11: معمارية نظام التوليد الآلي للنصوص

هذا الشكل هو تمثيل معماري للتوليد الآلي للنصوص، أبرز مكونتين فيه هما مصمم الخطاب ومحقق السطح. يقوم المصمم بتصفية المحتوى من قاعدة المعارف ويبنيه بحسب الغرض التواصلي. ثم يقوم محقق السطح بتوليد الجمل حسب الضوابط المعجمية والتركيبية التي زود بها النظام (برنامج التوليد) باتباع التصميم الذي خصصه المصمم[1].

[1] ينظر: Vander Linden (K.), Speech and language processing: An introduction to speech recognition, computational linguistics and Natural language processing, Prentice Hall, New Jersey, 1998, p. 6.

2. الأبحاث المنجزة

إن الأعمال الأولى، التي عالجت موضوع توليد النصوص بطريقة آلية، كانت تهدف إلى اختبار صلاحية نحو ما، كما فعل **فريدمان Friedman** حين أعد برنامجا خاصا يعين الدراسات التوليدية التحويلية. يعمل هذا البرنامج على إنتاج جمل إنتاجا عشوائيا فيختبر الباحثون بذلك قواعد النحو التي وضعوها لوصف اللغة وتفسير اشتغالها. لكن سرعان ما تخلوا عن هذا النوع من البحوث التي تهدف إلى اختبار نحوٍ ما، واستبدلوا بها البحوث التي تهتم بإنتاج جمل تحتوي على معلومة أو خبر. من هنا، فإن تحديد المحتوى الإعلامي المتعلق بالخطاب المرسل (أي ما الذي نقوله) ينتمي إلى مجال التوليد الآلي مثله مثل إنتاج الخطاب في ذاته (أي شكل ما نقوله).

إن المحتوى الإعلامي للرسالة يسبب مشكلا في تطبيقات المولد الآلي. ففي نظام الترجمة الآلية، ينبغي أن يكون المحتوى الإعلامي للنص المنتج مطابقا للمحتوى الإعلامي للنص المصدر الذي يعمل على ترجمته. وبالتالي، لا يعالج المولد المدمج بنظام الترجمة الآلية مشكل مضمون الخطاب بما أنه ثابت في النص المصدر وكذا في النص المستهدف أو المنتج[1].

بالمقابل، عندما يتعلق الأمر بإنتاج ملخص بطريقة آلية، يستخرج الباحث من النص الذي يريد تلخيصه المعلومات الضرورية لأنه بحاجة إليها حتى يثبتها ويضمنها النص الملخص. فقد ذكرت **دانلوص Danlos** أن أحد الباحثين قام بإعداد ملخصات لثلاث قصص تتضمن الحالات النفسية لبعض الأشخاص تعكس ثلاثة أوضاع نفسية إيجابية، ومحايدة، وسلبية. ثم قام برسم

---

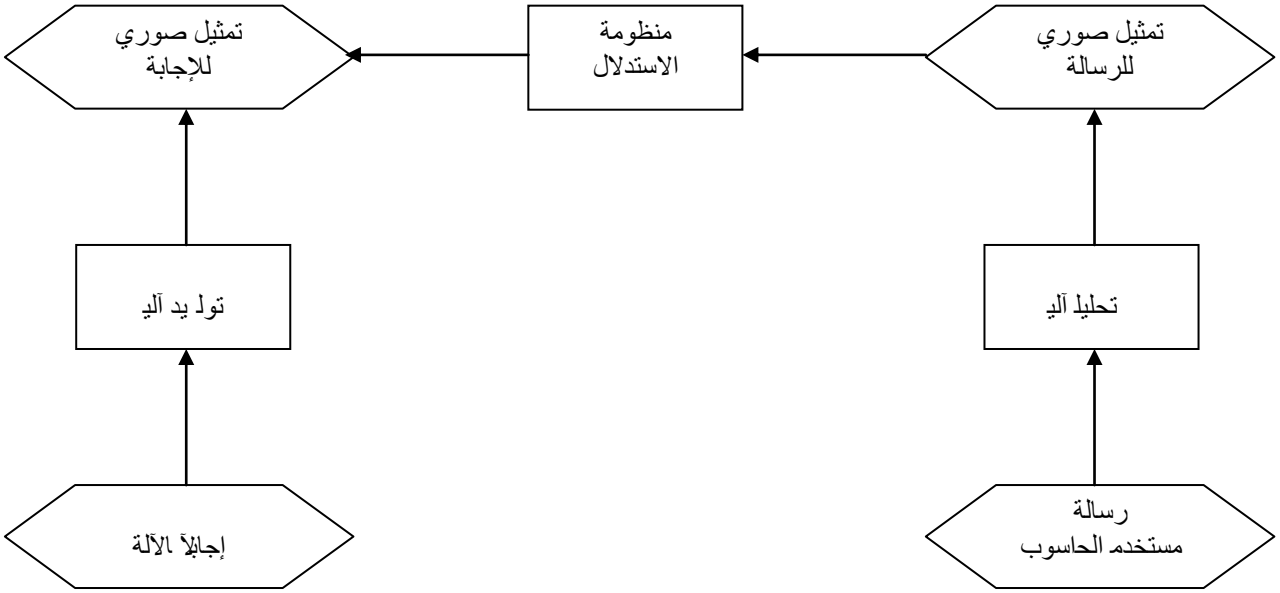
[1] ينظر: Danlos (Laurence), génération automatique de textes en langues naturelles, *op.cit*, p47.

بيان ذي عقد وفروع؛ تمثل العقد الأوضاع النفسية لأبطال تلك القصة، وتمثل الفروع العلاقات السببية بين تلك الأوضاع. بعد إقامة أصناف التكافؤ بين بيانات كهذه وتحقيق إجراءات خوارزمية تم الحصول على عقد خاصة تحدد مرتكزات القصة، وبالتالي شكلت من هذه المرتكزات عناصر الملخص.

يمكن للمولد أن يضاف إلى نظام خبير من أجل أن يعبر عن نتائج ذلك النظام الخبير باللغة الطبيعية، ويوفر تفسيرات لتلك النتائج. يتم توليد تلك التفسيرات عن طريق ترجمة قواعد البرنامج إلى اللغة الطبيعية بعدما نفذت للحصول على نتائج. أشارت **دانلوص Danlos** أيضا إلى أن هذه الطريقة قد انتقدت نظرا لأنها تعطي توضيحات حول اشتغال البرنامج ولا تقدم أي تبرير لذلك الاشتغال. ويرجع سبب المشكل حسب رأيها إلى أن المعارف الخاصة بكيفية تحقق البرنامج لا تنتمي بالضرورة إلى البرنامج نفسه.

إذن يرجع الفصل إلى الباحثين في مسألة: إن كانت الإحالة إلى الأشخاص والأشياء تدخل ضمن شكل الرسالة أو ضمن محتواها. تفترض **دانلوص Danlos** أنها تتناول هذا المشكل ضمن منظومتين مختلفتين: منظومة الاستدلال التي تعالج المحتوى الإعلامي للرسالة المنتجة، وخرج هذه المنظومة مترجم بواسطة منظومة أخرى مستقلة إلى اللغة الطبيعية، وقد بينت ذلك في الشكل الآتي[1]:

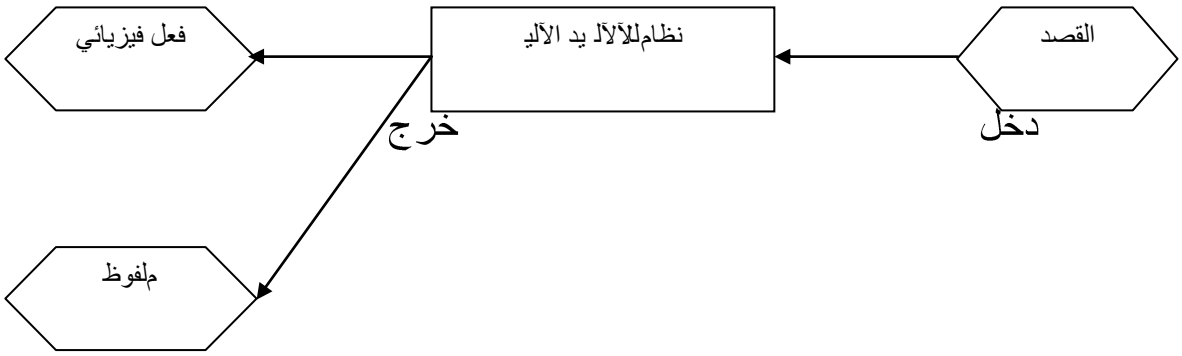
[1] ينظر: Danlos (Laurence), Génération automatique de textes en langues naturelles, *op.cit*, p. 11.



الشكل12: الحوار بين الإنسان والآلة

يمثل هذا الشكل كيف يتم التواصل بين الإنسان والآلة، وكيف يحصل تبادل المعلومة بينهما في ثلاث مراحل: المرحلة (1) يقوم فيها الحاسوب باستقبال رسالة المستخدم وتحليلها آليا ومن ثم يستخرج تمثيلها الصوري، ثم يقوم بتحديد كيفية الإجابة من خلال منظومة الاستدلال في المرحلة (2). وأخيرا يقوم في المرحلة (3) بصياغة الإجابة المناسبة انطلاقا من تمثيلها الصوري ويولدها في اللغة الطبيعية.

بالمقابل، هناك من الباحثين من يميل إلى الرأي بأن إنتاج الجمل هي وسيلة للفعل من أجل تحقيق هدف معين. هذا الرأي ينسجم مع الرؤية التداولية حيث تجعل صياغة الملفوظات مساواية للأفعال الفيزيائية في القدرة على التأثير في الواقع إذ يمكن أن أحرك كرسيا بنفسي أو أطلب من شخص فيحركه. إذن نظام التوليد الآلي حسب هذا الاتجاه مكون من دخل يمثل القصد والهدف من الرسالة، وخرج يتضمن الفعل الفيزيائي أو التلفظ بنص ما.



الشكل13: المولد الآلي من وجهة تداولية

إن نظام التوليد هذا يقوم على اختبار مختلف الحلول التي تؤدي إلى الهدف، ومختلف الوسائل التي تسمح بتحقيق تلك الحلول أيضا، كما يعتمد هذا البرنامج على نموذج الاستدلال المنطقي، وعلى قاعدة بيانات مزودة بأوصاف للموقف (حالة العالم) ولمعارف المتحاورين[1].

وعليه، فإن هذا الرأي الذي يتضمن جعل الملفوظات بمحاذاة الأفعال الفيزيائية في قوة التأثير يتعارض مع الرأي التقليدي الذي يميل إلى كون أفعال الكلام ترجمة للمعاني. لكن مع ذلك، تبقى الرؤية التقليدية صالحة دائما عندما لا تكون عملية إنتاج الكلام ضمن وضعية الحوار.

إن أنظمة التوليد الأولى لم تكن تطمح إلى توليد النصوص ضمن عملية الاتصال، بل كانت تقوم بتفريع جمل ذات محتوى إعلامي عشوائي لأن الهدف كان اختبار نظريات تركيبية، ثم ظهرت بعد ذلك أنظمة لا تستطيع أن تولد إلا جملا منعزلة، وعليه لم يتم تناول مشكل توليد النصوص ضمن عملية الاتصال إلا في حدود 1980.

[1] ينظر: Danlos (Laurence), Génération automatique de textes en langues naturelles, *op.cit*, p. 49.

في هذا الإطار، يجب أن يلبي النص المولد مطلبين هامين: من جهة ينبغي أن يحدد للمستخدم المعلومات التي يريدها، ومن جهة أخرى ينبغي أن يوفر صيغة لتلك المعلومات بلغة صحيحة. وبالتالي، فإن سيرورة التوليد تتضمن مؤلفتين: المؤلفة الأولى هي نظام خبير للاستدلال يعالج المحتوى الإعلامي للنصوص، والمؤلفة الثانية هي منظومة للتوليد اللساني تعالج صياغة ذلك المحتوى الإعلامي بلغة صحيحة[1].

3. مكونات المولد ووظائفها

إن مهمة نظام التوليد تتسم بالصعوبة والتعقيد وعدم التجانس، مما جعل الباحثين يقسمونه إلى عدة مكونات، يتكلف كل مكون منها بتأدية وظيفة معينة، وتتآزر تلك المكونات وتترابط من أجل توفير المعالجات المناسبة للتوليد.

1.3. مكونات المولد

إن عملية توليد النصوص مثلها مثل كل العمليات المعقدة يمكن أن تتشعب إلى مجموعة من العمليات الجزئية. وعليه، فإن مولد النصوص يتكون على الأقل من أربعة عناصر –كما هو مبين في الشكل14– هي:

- **المصمِّم الكلي:** هو منظومة مفهومية مهمتها إعطاء الأهداف والموارد المعرفية (المعطيات الخام) لتعيين الرسائل والبنية العامة للنص، ينتج عن ذلك عموما مشجر مؤلف من منطلقات هي الرسائل التي سيعبر عنها ومجموعة من العقد والأسهم التي

[1] ينظر: Fuchs (C.) et *al*., Linguistique et traitement automatique des langues, Hachette, Paris, 1993, pp. 248-251.

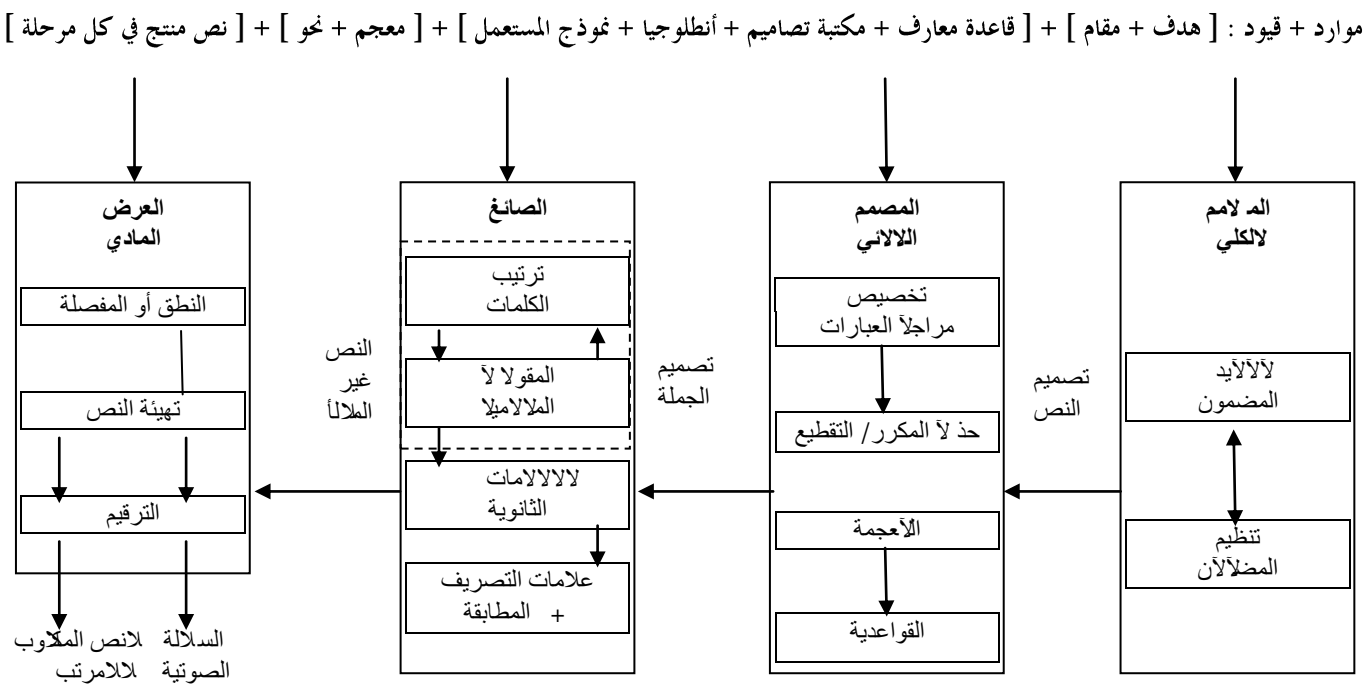
تعبر عن نوع العلاقة الخطابية (النتيجة، السببية...)، والوضعية النصية للمقاطع أو السياق؛

- **المصمِّم الجزئي:** يقع بين المنظومتين المفهومية واللسانية، يقرر كيفية وصف واقعة أو كيان أي يقدم معلومات كافية تسمح بتمييز الوقائع والكيانات المختارة ضمن مجموعة من الوقائع والكيانات الممكنة ومراجعها؛ كما يقوم بجمع المعطيات المتقاربة مع إسقاط المكرر منها من أجل إنشاء نص مختصر، عدة وسائل يمكن استخدمها من أجل ذلك منها الإضمار؛ اختيار الكلمات المناسبة للتعبير عن المضمون المختار ومن أجل إتمام درجة الاتساق (المعجمة[1]، الإضمار، اختيار الكلمات الثانوية). من الناحية التطبيقية، لا يمكن الفصل بوضوح بين اختيار العناصر المعجمية وبين الالتزامات التنظيمية للنحو.
- **الصائغ أو محقق السطح:** هو المكون اللساني لهذه العملية، تنحصر مهمته في وضع الكلمات في أماكنها أي ترتيبها خطيا كما يفرضه النحو ووضعية المعلومة كما تتطلبها اللغة، وتعيين مقولاتها المعجمية (اسم، فعل، صفة...)، وإدراج الكلمات الثانوية (حروف الجر)، والقيام بالعمليات الصرفية التركيبية (إلحاق علامات التصريف، المطابقة)؛
- **المؤلِّف:** هو منظومة تتضمن تهيئة النص في شكله المنطوق أو المكتوب، وترجمة سلسلة الحروف إلى شكلها الصوتي أي إلى فونيمات، وإعطاءها التنغيم المناسب.

---

[1] **المعجمة *lexicalisation*** عملية تحويل المفاهيم إلى لوسيمات مثلا: المفهوم **LOVE** يمكن أن يعبر عنه في العربية بالاسم "**حب**" أو الفعل "**أحب**"...

إذن تتكون عملية التوليد من مهام ينجزها كل من المصمم الكلي، المصمم الجزئي، الصائغ، وتحتاج في ذلك إلى موارد وتلتزم بقيود، كما هو مبين في الشكل[1].



الشكل14: مكونات المولد ووظائفها

2.3. وظائف مكونات المولد

يتألف المولد من عدة مكونات تؤدى كل واحدة منها أدوارا ووظائف معقدة ومتداخلة من أجل توليد النصوص والجمل.

1.2.3. التصميم الكلي

هناك عدة مقاربات مقترحة في تصميم المحتوى، وهي تختلف في نقطة بداية التوليد العميق بعضها يعتقد بأن محتوى الرسالة يأتي من منظومة أخرى خارجية عن منظومة التوليد،

[1] ينظر: Mitkov (R.), The Oxford Handbook of computational linguistics, *op.cit*, pp. 288-289.

وبعضها يوظف المخططات ونظرية البنية الخطابية من أجل تحديد محتوى الرسالة. تتولى منظومة التصميم الكلي، التي تدعى مصمم النص أو المولد العميق، مهمتي اختيار المضمون العميق والبنينة الخطابية.

- **اختيار المضمون العميق**

الدخل	الخرج
– قواعد معارف حاسوبية – أهداف اتصالية	– شبكات مفهومية – شبكات دلالية

أثناء اختيار المضمون العميق، يقوم المولد بتخصيص المحتوى الدلالي للنص المراد إنتاجه. وبالتالي، يشغل عمليات اختيار المعلومات ونسخها في شبكات أو تمثيلات دلالية، مفهومية، منطقية... وبما أن المولدات تتعدد، والمعطيات التي تشغلها كذلك، فإن الطرق المستخدمة ستكون متعددة هي أيضا، وعليه يميز الباحثون ثلاث مقاربات مختلفة لمعالجة هذه العمليات:

ففي بعض المولدات يتولى المستخدم عملية اختيار المحتوى العميق إما بصفة جزئية أو كلية، وذلك بوساطة نظام تصافح يسمح بإدخال مضمون، وأحيانا يتولى هذه المهمة تطبيق خارجي عن المولد (نظام خبير، مولد مهام...) هو الذي يتولى عملية اختيار المحتوى العميق. في هاتين الحالتين، تنحصر مهمة المولد في ترجمة المعطيات التي تم اختيارها إما من المستخدم أو من التطبيق الخارجي عنه إلى بنى دلالية أو مفهومية تحضيرا للبناء اللساني. يتم

غالبا في هذا النوع من المقاربات اختيار المحتوى العميق بمعية البنينة الخطابية التي تتولى إنشاء البنية البلاغية ذات العلاقات المنطقية للنص.

نوع آخر من المولدات يأخذ في المدخل الأهداف الاتصالية بحيث تتم نمذجة الهدف الاتصالي عموما على أنه حالة ذهنية للمستخدم يريد أن يصل إليها المولد عن طريق فعل اللغة. هذه الحالات تمثل في أغلب الأحيان معرفة أو كفاية. وعليه، يتم اختيار المضمون العميق عن طريق تبسيط الأهداف الاتصالية إلى أهدف جزئية، ويبحث المولد عن الهدف الاتصالي المزود به والموافق للهدف الاتصالي الجزئي. كل هدف يقدم مكملا للمعلومة ويحول بعد ذلك إلى فعل اللغة (مثلا: إخبار، توصية، أمر، تفسير تحول كلها إلى جملة).

النوع الثالث من المولدات ينطلق من رصيد بيانات أو معارف تابعة لميدان معين. في هذه المقاربة أيضا، يمكن للنظام أن يتناول أهدافا اتصالية، لكنه لا يقود عملية اختيار المضمون بالطريقة نفسها. في القاعدة العامة، يتعلق الأمر بأهداف أكثر عمومية وضبابية، مثل وصف موضوع أو واقعة. ومحتوى الرسالة إذن هو مجموعة جزئية لقاعدة المعارف. ينبغي –ضمن هذه المقاربة– الارتكاز على مخططات، قواعد تجريبية... في عملية اختيار المضمون.

- **البنينة الخطابية**

الدخل	الخرج
– شبكات مفهومية – شبكات دلالية	– تصميم النص

تشير البنية الدلالية المتحصل عليها خلال عملية اختيار المضمون العميق إلى المعنى العام للنص، لكن يبدو أنها تحتاج إلى معالجة إضافية قبل أن تحال على المكونات اللسانية. بداية عندما يتضمن المضمون العميق للرسالة عددا كبيرا من المعلومات، ويبلغ النص حجم فقرة أو فقرات، تصير البنية الدلالية موضوعا معقدا ومكثفا تضع المنظومتين التركيبية والمعجمية في مواجهة توليفة من الاختيارات التي لا تخدم المعالجة. وعليه، فإن البنية الدلالية لا تأخذ في الاعتبار المسار الخطابي أي ترتيب الأفكار والعلاقات التي تربطها فيما بينها. وبالتالي، تدعو الحاجة إلى إنشاء بنية خطابية من أجل تنظيم الرسالة في بناء متناسق. في بعض الأنظمة، تدمج هاتان العمليتان (اختيار المضمون العميق والبنينة الخطابية) في منظومة واحدة، وتفضل بعضها أن تجعل البنينة الخطابية تالية لعملية اختيار المضمون العميق.

من الناحية التقنية، تتضمن البنينة الخطابية تقطيع المحتوى الدلالي إلى وحدات تتحقق بعد ذلك في شكل جمل أو لفظات، كما تتضمن مفصلة تلك الوحدات في تصميم نص متناسق. ترتبط تلك الوحدات بواسطة علاقات خطابية، ويستعمل لذلك عادة نموذج البنية الخطابية، تقوم تلك العلاقات الخطابية من جهة بالربط بين الجمل واللفظات، ومن جهة أخرى بتوضيح العلاقات التداولية والخطابية مثل التفسير، النتيجة، البلورة، الهدف... وجود علاقة خطابية يسمح باختيار رابط خطابي (حرف عطف، رابط تبعية: موصول حرفي أو اسمي...). تمثل العلاقات الخطابية مفصِلا بين العلاقات الدلالية والبناءات التركيبية.

العلاقات الخطابية هي قريبة من العلاقات الدلالية، وهي لا تربط بين الوحدات الصغرى للمعنى بل بالأحرى بين تجمعات معقدة للمعنى، وهي تعبر عن نحو النص. يتم اختيارها عموما

بالاستنتاج انطلاقا من علاقة دلالية أو مفهومية، لكن إسقاط المستوى الدلالي على المستوى الخطابي ليس وحيد الاتجاه أو مبتذلا. مثلا: يمكن لعلاقة دلالية مسببية بين عملية ووضعية أن تعطي سلسلات خطابية مختلفة:

– من أجل إضاءة الغرفة، اضغط على الزر. (الهدف)

– الضغط على الزر يضيء الغرفة. (الأداة)

2.2.3. التصميم التركيبي

الدخل	الخرج
– تصميم النص – شبكات دلالية	– مشجرات تركيبية

ضمن هذه المهمة، يباشر النظام اختيار البناءات النحوية التي ستصمم جمل النص. يتعلق الأمر هنا بتحويل البنية الدلالية مع احتمال تزويدها ببنية خطابية إلى مشجر تركيبي، حيث يتم إسقاط العلاقات الدلالية على العلاقات أو الوظائف التركيبية. مثلا: العلاقة الدلالية (السببية) يمكن أن تعطي في العربية اللام: "انهزم المصارع لعدم تحضيره الجيد". هناك عدد من أنظمة التوليد تستعمل معاجم–أنحاء مدمجة أو أنحاء معجمية حيث تكون البناءات التركيبية مثبتة بمداخل معجمية، وبالتالي يتم التصميم التركيبي بالموازاة مع المعجمة.

3.2.3. المعجمة

الدخل	الخرج
– شبكات دلالية – مشجرات تركيبية عميقة	– شبكات معجمية – مشجرات تركيبية سطحية

المعجمة هي المهمة المكملة للتصميم التركيبي بما أنها تتضمن اختيار كلمات اللغة التي يجب أن تجسد المعاني المحددة في المراحل السابقة. هذا الاختيار يمس فقط الكلمات المليئة (اسم، فعل، صفة...) وليس الكلمات الأدوات (حروف الجر، التعريف، العطف...) التي تنتمي إلى التصميم التركيبي. كما سبقت الإشارة إليه، ترتبط المعجمة بالتصميم التركيبي، ففي الواقع تحويل وحدة معنوية أو مفهومية إلى عنصر معجمي يستدعي اختيار مقولة نحوية له (اسم، فعل، صفة...).

من الضروري هنا الحرص على إيجاد الملاءمة بين الكلمات والمركبات وبين مقولاتها النحوية. مثلا: المركب "إصلاح الجهاز" يمكن أن يكون معمولا لحرف الجر (قصدت محلا لإصلاح الجهاز)، دون المركب "أصلح الجهاز" المرادف له دلاليا الذي لا يمكن أن يحل محله، فيكون معمولا لحرف الجر.

4.2.3. الإحكام الصرفي

الدخل	الخرج
– مشجرات تركيبية	– مشجرات تركيبية موسعة – سلاسل خطية إملائية

بعد مرورها بالتصميم التركيبي والمعجمة وفي سياق عملية التركيب، تُعرض الرسالة في شكل مشجر تركيبي تحتل عقده عناصر معجمية أي مداخل قاموسية. ومن أجل الحصول على الشكل السطحي والصحيح للنص، تتم التسوية الخطية لمشجرات المكونات، ويبقى تصريف الكلمات الواردة من قاموس التطبيق: تصريف الأفعال، مطابقة الأسماء والضمائر... وأيضا الظواهر التي تحدث لبعض السلاسل عند إلصاق أجزائها كما هي الحال في اللغة الفرنسية (*de +le = du*)، وفي اللغة العربية يمكن أن نذكر كمثال: حذف النون خطّا عند إدغامها في الميم (مِنْ + مَنْ مِمَّنْ) أو إبدال الألف المقصورة في "إِلَى" ألفا ممدودة عند إلصاقها بـــ"مَا" الاستفهامية مع حذف ألفها "إِلاَمَ"...

5.2.3. التهيئة الطباعية

الدخل	الخرج
– مشجرات تركيبية سطحية – سلاسل حروف مكتوبة	– ملف النص المهيأ للطبع

التهيئة الطباعية عملية ضرورية لجعل النص مقروءا للمستخدم، وهي مرحلة كمال العمل وإتمامه حيث يتم فيها إضافة علامات الترقيم، وإجراء بعض التعديلات كتكبير الحرف الأول من الكلمة المبدوء بها أو من اسم العلم في اللغات ذات النظام الكتابي اللاتيني، وكوضع الفراغات بين الكلمات والفقرات. تنتج أنظمة التوليد نصوصا مشفرة في اللغات الحاسوبية الخاصة بتأليف النصوص مثل ***SGML/XML, HTML, LaTeX***[1].

4. نماذج من أنظمة التوليد المنجزة

المولدات هي أنظمة ترتبط غالبا بميدان وتطبيق معينين؛ فالنظام هو « مجموعة متناسقة من الممارسات أو التصرفات، تُنجز بهدف الحصول على نتيجة »[2]، وبالتالي فإن أنظمة التوليد مكونة من مجموعة من المنظومات (أنظمة جزئية) تقدم نتيجة لاشتغالها نصوصا، ويمكن أيضا لنظام التوليد أن يشكل منظومة لنظام أكبر (كنظام الحوار بين الإنسان والآلة...). والميدان هو « الذي يحتضن فنا أو علما أو موضوعا أو فكرة »[3]، لذلك هناك ميادين عامة ترتبط بها المولدات: طبية واقتصادية... وأخرى أكثر حصرا كميدان المناخ البحري والبورصة... والتطبيق هو « عملية الاستناد على شيء، وهو في المعلوميات: برنامج مكتوب من أجل استعمال دقيق »[3]، وبالتالي يمكن تعريف مولد المغرابي مثلا –والذي سيأتي الحديث عنه– بأنه نظام إنتاج وصفات للمطبخ باللغة العربية، أو تطبيق لإنتاج النصوص العربية في ميدان وصفات المطبخ.

---

[1] ينظر: Pierrel (J.M.), Ingénierie de la langue, Hermès, 2000, pp. 311-316

[2] ينظر: Rey-Debove (Josette) & Rey (Alain), Le nouveau Petit Robert, Dictionnaires Le Robert, 1993, p. 673, p. 104.

يعتمد الباحثون في دراسة أنظمة التوليد على مجموعة من الخصائص، وتستعمل تلك الأنظمة غالبا اللغات التي لها حضور ثقافي واقتصادي في العالم، وتسعى إلى تحقيق مجموعة من الأهداف التي من أجلها قام الباحثون بإنشائها. أحاول أن أعرض في الجدول الموالي نماذج من أنظمة التوليد التي يعتمد عليها الباحثون في توليدات مختلف اللغات[1].

[1] للاطلاع على أنظمة التوليد الآلية يرجى الرجوع إلى الصفحة الإلكترونية على العنوان: http://purl.org/net/nlg-list حيث يقوم المشرفون على الموقع في كل مرة بإضافة مولدات مختلف اللغات الأوروبية بالإضافة إلى اللغات الصينية واليابانية والعبرية... ولم ترد في القائمة مولدات اللغة العربية !

النظام	اللغة	الهدف	ماذا يتناول ؟	العمليات اللسانية	المداخل	التمثيل الداخلي	اختيار الكلمات
BABEL	الإنجليزية	توليد الجمل، والجمل التي لها نفس المعنى	كيف يقال ؟	؟	مفاهيم	بنية تعلقات مفهومية	معجم مفاهيم
Simon & Slocum	الإنجليزية	تحويل بنية عميقة إلى لغة طبيعية	ماذا يقال ؟ كيف يقال ؟	؟	شبكات دلالية	ATN	؟
BLAH	الإنجليزية	إنتاج إيضاحات	كيف يقال ؟ خاصة	العمليات: و/أو إذا/فـــ	؟	مشجر	؟
XPLAIN	الإنجليزية	إنتاج إيضاحات	كيف يقال ؟ ماذا يقال ؟	عمليات	؟	نظام خبير	؟

Danlos	الإنجليزية/الفرنسية	توليد برقيات صحفية	كيف يقال ؟	عمليات	استفسار	لغة وسيطة	متوفر
Simonin	الفرنسية	توليد وبناء نص انطلاقا من تمثيل داخلي	ماذا يقال ؟	عمليات	؟	أوصاف وظيفية	متوفر
Boyer et Laplame	الفرنسية	توليد كل الجمل المترادفة الممكنة لشبكة دلالية	كيف يقال ؟	؟	تعيينات مفهومية	شبكات دلالية	متوفر
Nogier	الفرنسية	نظام أسئلة/إجابات	كيف يقال ؟	؟	كلمات	بيانات مفهومية	متوفر

الجدول7: نماذج لأنظمة التوليد

على اعتبار أن عملية التوليد تتكون من مؤلفتين استراتيجية خاصة بمحتوى القول (ماذا يقال ؟) وتكتيكية لها صلة بطرق القول (كيف يقال ؟)، فإن بعض المولدات تحتوي على كلتا المؤلفتين كما هي الحال في نظامي **Simon & Slocum** و**XPLAIN**، أو تحتوي فقط على إحداهما مثل نظام **BABEL** الذي يتناول المؤلفة التكتيكية، وكذا نظام **Simonin** الذي يقتصر على المؤلفة الاستراتيجية. كما أن هاته الأنظمة تجري مجموعة من العمليات تتمثل في تحويلات لسانية تغير من تمثيل الرسالة دون المساس بمحتواها الدلالي مثل نظام **BLAH** الذي يقوم بوصل الجمل بواسطة روابط منطقية (و/أو)، (إذا/فـــ).

بالنسبة إلى مداخل أنظمة التوليد، تتعدد بين كونها مفاهيم، أو شبكات دلالية، أو كلمات، أو تكون على شكل استفسار في نظام **Danlos** يقوم بملئه المستخدم. وتتعدد أيضا أدوات التمثيل داخل المولدات فقد تكون على شكل بنيات تعلق أو مشجرات أو شبكات دلالية، أو تستخدم فيها تقنيات ***ATN*** (شبكة الانتقال المعززة) التي تسمح بالتسوية الخطية للبنية التركيبية التراتبية، وتستخدم في بعضها صورانيات اللغة الوسيطة أو نظام خبير أو بيانات مفهومية، هذه الأخيرة التي تعتمد عليها المؤلفة الاستراتيجية في كثير من أنظمة التوليد لأنها تسمح بتمثيل دلالي غير مرتب على عكس منطق المحمولات الذي

يسمح من جهته باستنتاج عدة بنى لسانية انطلاقا من بنية معنى معين[1]. في الأخير، معظم الأنظمة الواردة في الجدول تتمتع بمنظومة لاختيار الكلمات.

بالإضافة إلى ذلك، هناك بعض البرامج مثلا: ***HAIKU, RENGA, ROMAN*** أعدت لتولد أشعارا ونصوصا أدبية قصيرة بأسلوب قريب من أسلوب المؤلفين السرياليين. يتم التوليد انطلاقا من نص نوعي مكون من جمل ذات فراغات (قوالب)، كل فراغ يحتوي على واصف يسمح بتصفية الكلمات التي يمكن أن تشغله. هذه الواصفات الرقمية تتضمن المعلومات الدلالية والتركيبية المتنوعة مشفرة، كما تأخذ بعين الاعتبار الظواهر الخاصة بالنصوص الشعرية[2].

1.4. نموذج من مولدات اللغة العربية

على عكس ما يمكن توقعه، لقد نال اللسان العربي نصيبه من اعتناء الباحثين في إنشاء مولدات للنصوص العربية حتى قبل ألسنة أخرى أوروبية كالألمانية والإيطالية، وذلك منذ عهد مبكر في سنوات 1980، لذلك سأتعرض لإحداها والتي قامت بإنشائها الباحثة شادية المغرابي.

---

[1] **البيانات المفهومية *graphes conceptuelles*** هي لغة لتمثيل المعارف في مجال الذكاء الاصطناعي، وهي تقع في مسار الأنطولوجيات الصورية والشبكات الدلالية. وهي ترتكز على نظام ترميز بياني ترتبط فيه عقد مفهومية مع عقد علائقية. ينظر: Pierrel (J.M.), Ingénierie de la langue, *op.cit*, p.312

[2] ينظر: Balicco (Laurence), *Génération de répliques en Français dans une interface Homme-Machine en langue naturelle*, thèse de doctorat, Université de Grenoble, 1993, pp. 22-44.

الخصائص	
اللغة	العربية
الميدان	وصفات للمطبخ
الخرج	النص

عثرت في أحد المراجع على معلومات مقتضبة حول هذا المولد؛ فهو يقوم بإنتاج وصفات للمطبخ باللغة العربية. تمثل الوصفة عن طريق أفعال بسيطة متتابعة ينجزها إنسان آلي. وقد ابتكر هذا التمثيل خصيصا لهذا الميدان. يستعمل مولد المغرابي مجموعة من القواعد لاختيار كلمات النص، وهي أيضا مصممة لهذا الميدان؛ مثلا: هناك زهاء 20 قاعدة مرتبطة باختيار كلمة للتعبير عن المفهوم "انتظار"، وهي تأخذ في الاعتبار الوضعية الحالية والوضعية المرجوة للجزء الذي تتم معالجته بحسب ما إذا بلغ المولد نتيجة سلبية أو بلغها عبر سلسلة من الأفعال. مثلا:

إذا كانت درجة الحرارة الحالية: T1

إذا وجب انتظار بلوغ درجة الحرارة: T2

إذا كانت $T1 < T2$

إذا كانت T2 مساوية للدرجة الحرارة المحيطة

إذن استعمل: اتركه يبرد.

عندما تكون هناك عدة كلمات متاحة للاستعمال، يفضل النظام تلك التي تعبر عن عدة أفعال في نفس الوقت. في المستوى اللساني، يحتوي المولد على مجموعة أخرى من القواعد من أجل تحديد الوضعيات التي يمكن أن تستخدم فيها الضمائر. ويملك المولد أخيرا مجموعة من القواعد التي تسمح بتحسين النصوص (مثلا تحديد نفس الفعلين في جملتين متواليتين، تفضيل المواضيع المشتركة...)[1].

2.4. ملاحظات حول أنظمة التوليد

تعمل الأنظمة المتوفرة في اتجاه خطي حيث تنتقل من التمثيل الدلالي إلى الشكل السطحي للنص، فتتوالى مختلف الوضعيات في المعالجة في اتجاه واحد، ويتم الانتقال من منظومة لأخرى مع عدم إمكانية العودة إلى الوراء. هذه المعمارية المنظومية الخطية تمنح للنظام قدرة على التطور والصيانة حيث يتسنى للباحثين تدارك النقائص الموجودة في منظومة من المنظومات التي يتكون منها المولد خلال مرحلة من مراحل المعالجة، واقتراح بدائل لمواطن النقص والخلل.

من بين النقائص التي لوحظت على أنظمة التوليد المتوفرة، ارتباطها الوثيق بتطبيق خاص وغير ملائمتها لتطبيقات أخرى. كما أنها غالبا ما تكون متصلة بلغة معينة قد يصعب تعميمها على لغات أخرى، وتتميز النصوص التي تنتجها بشكلها الجامد. بالإضافة إلى ذلك، لا يرتكز الجانب اللساني للأنظمة المدروسة على نموذج لساني إلا

---

[1] ينظر: Ponton (Claude), *Génération automatique de textes en langue naturelle : essai de définition d'un système noyau*, thèse de doctorat, Grenoble 3, 1996, p. 248.

جزئيا، فقد يقوم النموذج اللساني على مجموعة من القواعد النحوية تكون مدرسية في العادة، ولا يتسم بالصرامة الكافية والكاملة حتى يغطي حالات عديدة ولغات مختلفة.

ويلاحظ أيضا أن تلك الأنظمة ليست مفتوحة أي سهلة التغيير والإضافة، فحينما يكتمل وضعها يتعذر أو يستحيل تحيينها بإضافة أشياء أخرى عليها. ومن الناحية التقنية، لا يمكن نقل تلك الأنظمة وتعميمها على كل أصناف الآلات (عدة عوامل تدخل هنا كسعة الذاكرة وسرعة المعالجة)[1]. كل هذه النقائص لا تساعد كثيرا على إعادة استعمال تلك الأنظمة، والقليل منها يستخدم اللغة الفرنسية فضلا عن اللغة العربية التي تبقى بحاجة إلى المصالحة مع التقنية والعلوم الحديثة وإلى بذل المختصين العرب الجهد في البحث وصدقهم في ذلك، على عكس اللغة الإنجليزية التي تستخدم فيها المولدات الآلية بكثرة.

5. تقنيات التوليد

لتحويل معطيات غير لسانية إلى لغة طبيعية، تتضمن سيرورة التوليد ثلاث مراحل ممثلة في الجدول الآتي:

---

[1] ينظر: Balicco (Laurence), *Génération de répliques en Français dans une interface Homme-Machine en langue naturelle*, *op.cit*, p. 61.

	ثلاث مراحل لسيرورة التوليد
(أ)	تعيين الأهداف المراد تحقيقها في الجملة
(ب)	تصميم الأهداف
(ت-1)	إنشاء التصاميم
(ت-2)	تحقيق التصاميم في نص

- **تعيين الأهداف**

من المعلوم أن التوليد وكذا كل نشاط الإنتاج الكلامي تحفزه أهداف لتلبية طلبات معينة. مثلا عندما أريد أن يحقق أحد ما الهدف: إغلاق النافذة لأنني أشعر بالبرد، فإنني أطلب في صورة خبر أو استفهام مثلا:

أ- البرد قارس.

ب- هل يمكن أن تغلق النافذة ؟

يمكن تمثيل ذلك في العبارة التالية: **طلب(س، إغلاق(ع،النافذة))**

حيث **س** هو فاعل الهدف **(طلب)** والمحمول **إغلاق(ع،النافذة)** يدل على الفعل الذي يقوم به **ع**.

- **تصميم الأهداف**

بعد تعيين الأهداف أو المضمون، ينبغي على المولد أن يصممها، أي أن ينظمها في تصميم أو مخطط تبعا للمعلومات الأكثر توافقا مع غرض التوليد.

يسرد المخطط إذن سلسلة الأفعال التي يجب أن يباشرها المولد لتلبية الأهداف. يصف دلالة النص من خلال تخصيص مختلف الجمل التي يراد توليدها[1].

اعتمادا على المعلومات اللسانية الممثلة فيما يلي يمكن توليد عدة جمل:

ميلاد	فاعل	الأم
	موضوع	الابن
الابن	[اسمه	محمد]
المكان	[الجغرافيا	البلد]
البلد	[اسمه	مكة]
الزمان	[التاريخ	السنة]
السنة	[اسمها	عام الفيل]
نشر	فاعل	الإنسان
	موضوع	الإسلام

إذن يمكن أن أولد الجملتين الآتيتين من بين جمل أخرى ممكنة:

[1] ينظر: Bouillon (P.) et *al*., Traitement automatique des langues naturelles*, op.cit*, p. 183

- الجملة (1): ولد محمد (صلى الله عليه وسلم) بمكة عام الفيل، ونشر الإسلام.
- الجملة (2): كان ميلاد محمد (صلى الله عليه وسلم) الذي نشر الإسلام بمكة.

للحصول على الجملتين، أقوم بوضع مخططين لكل منهما من خلال الجدول الموالي:

	تصميم الأهداف
مخطط_1	م_1 (نوع التصميم–الجملة ج_1) (القضية ق1 م1،م2)) (م1 (محمول = ميلاد) (متعلق1 = o) (متعلق 2 = محمد)) (القضية ق2 م1،م2)) (م2 (محمول = نشر) (متعلق1 = محمد) (متعلق 2 = الإسلام))
مخطط_2	م_2 (نوع التصميم–الجملة ج_2) (القضية ق2 تابعة للقضية ق1) (القضية ق1 م2) (القضية ق2 م1،م2))

يبين هذا الجدول العناصر التالية:

- مخطط_1 ومخطط_2: يعينان مختلف التصاميم الممكنة؛
- م_1 وم_2: يشيران إلى التصميمين الفرعيين المتصلين على الترتيب بـــ مخطط_1 ومخطط_2؛
- ج_1 وج_2: يدلان على الجملتين المتصلتين بـــ م_1 وم_2؛
- ق1 وق2: يعينان القضايا المتصلة بالجملتين على التوالي؛
- م1 وم2: يشيران إلى المحمولين المرتبطين بالجملتين. كلا المحمولين لهما دلالة خاصة به على شكل بنية: محمول–متعلق.

إذن يشير مخطط_1 إلى أن النص المولد سيتكون من قضيتين ق1 وق2 بينهما علاقة عطف كل منهما يقدم معلومات جديدة وإضافية، وأن مخطط_2 الذي يحتوي على قضيتين ق1 وق2 مع كون ق2 تابعة للقضية ق1، بعبارة أخرى ق2 هنا لا تقدم معلومات إضافية[1].

- **إنشاء المخططات وتحقيقها**

إذا كانت مرحلة التصميم تسمح بتحديد المخططات الممكنة، فإن مرحلة صياغة المخططات في نص (إنشاء التصاميم (ت–1)) ستقوم باختيار أفضل مخطط وفقا لمختلف المعلومات التداولية: ما هي الصلات الخطابية (البلاغية) التي تجمع بين القضايا (السببية،

[1] ينظر: Bouillon (P.) et *al*., Traitement automatique des langues naturelles, *op.cit*, p. 185

التحفيز)، ما هو التيم الخاص بالمخطط (عن ماذا نتحدث ؟ موضوع النص–المخبر عنه)، ما هو الريم الخاص بالرسالة (ماذا نقول من جديد عن التيم–المخبر به)...

وبالتالي، إذا كنت أريد تقديم معلومات جديدة عن محمولي ق1 وق2 فإني أستخدم مخطط–1 للجملة (1)، أما إذا كنت أريد التركيز على القضية ق1 فإني أستخدم مخطط_2 للجملة (2) لأن ق2 هنا تابعة لــق1.

وفي الأخير، في مرحلة تحقيق التصاميم، تحول المخططات إلى نص في اللغة الطبيعية. تتضمن هذه المرحلة على الأقل التعبير عن الكلمات وعن محيطاتها التركيبية في السياق اللساني، وذلك بمراقبة ملحقاتها الصرفية وعلامات الترقيم في النص. مثلا يمكن أن يتحقق مخطط_1 في الجملة (1) المذكورة.

تتكفل بهذه المراحل الثلاث الضرورية عادة المؤلفتان الاستراتيجية والتكتيكية. المؤلفة الاستراتيجية تنتمي أساسا إلى العلوم المعرفية، والمؤلفة التكتيكية إلى اللسانيات الحاسوبية[1].

---

[1] ينظر: Bouillon (Pierrette) et *al*., Traitement automatique des langues naturelles, *op.cit*, p.186.

سيرورة التوليد وفقا للمؤلفتين الاستراتيجية والتكتيكية	
لماذا نقول ؟	(أ) التعبير المستهدف
ماذا نقول ؟ (المؤلفة الاستراتيجية)	(ب) تصميم الأهداف (أو اختيار مضمون النص المولد)
كيف نقول ؟ (المؤلفة التكتيكية)	(ت-1) إنشاء التصاميم واختيار الكلمات (أو تنظيم المضمون)
	(ت-2) اختيار البنى التركيبية المرتبطة بها؛ الملحقات الصرفية، علامات الترقيم

6. خوارزمات التوليد

تسمح الخوارزمات بوصف المخطط الذي يتبعه الحاسوب في معالجته لمختلف البيانات، وبما أن عملية توليد النصوص تتم معالجتها آليا، فإنها ترتكز على مخططات في إنجازها. هذه العملية التي تنطلق من تمثيل دلالي، فتعطي المشجرات التركيبية للجملة بالاعتماد على قواعد إعادة الكتابة. وعليه يتبع الباحثون في وصف عملية توليد النصوص ثلاث استراتيجيات هي:

- **الاستراتيجية التصاعدية**

وهي موجهة بواسطة المعلومات اللسانية النحوية والمعجمية، إذ تنطلق من الرمز الابتدائي للنحو، وتحاول بناء المشجر التركيبي من الأسفل إلى الأعلى من قمة المشجر إلى الأوراق أي كلمات الجملة المراد تحليلها.

- **الاستراتيجية التنازلية**

وهي موجهة بواسطة كلمات الجملة حيث تأخذها كنقطة لانطلاق التحليل، وتحاول بناء المشجر إلى غاية القمة، على عكس الاستراتيجية المتقدمة، فإن عملية التحليل تبدأ من النتائج نحو العناصر المجردة وصولا إلى العنصر الابتدائي[1].

يتميز التوليد الآلي للجمل من بين مجالات المعالجة الآلية للغات الطبيعية بأنه مجال بحث ثري وخصب، والمستخدمون في أمسّ الحاجة إلى تطبيقاته المتعددة، وهو يحاول أن يطور آليات جديدة لمعالجة عملية التوليد، وأن يتجاوز مختلف الصعوبات الناجمة عن اختيار عناصر معينة من بين عناصر ممكنة عند الانتقال من مستوى لساني لآخر.

وعليه، يشكل التوليد الآلي للجمل العربية رهانا لمختلف البرمجيات التي تعمل على استثمار المعارف اللسانية العربية، ومنها المعارف الخاصة بتكوين الجمل في تطبيقات تتناول خدمات متنوعة كتقارير الأرصاد الجوية، وملخصات النصوص وغيرها، وهي تحاول في ذلك أن تستفيد من مناهج اللسانيات الحاسوبية الموظفة في هذا الغرض، والتي قد أثبتت جدارتها في لغات أخرى.

---

[1] ينظر: lyse et de génération pour la aa eee rr rr rrrr rrr rrr rrrr rrrrr rrrrr r langue naturelle, Ed. Adosa, 1994, p. 54.

الفصل الخامس:

التوليد الآلي للجمل العربية

الفصل الخامس: التوليد الآلي للجمل العربية

في الفصول المتقدمة، تعرضت لأبرز الجوانب النظرية التي تشمل موضوع هذه الرسالة والمتمثلة أساسا في مفهوم الجملة قديما وحديثا، ومفهوم توليد الجمل الذي تتبعناه عند مختلف النماذج اللسانية. وبعد أن قدمت لسيرورة التوليد من وجهة نظر اللسانيات النفسية، عرجت أخيرا على التوليد الآلي للنصوص حيث وقفت على أبرز التقنيات المستخدمة في التوليد، وعرضت بعض المولدات المنجزة.

وقد سبقت الإشارة إلى أن نظام التوليد يتكون من مؤلفتين استراتيجية وتكتيكية بحيث تتكلف الأولى بتحديد محتوى النصوص، ومهمة دراسة هذه المؤلفة ترجع إلى العلوم المعرفية، أما المؤلفة التكتيكية فمهمتها منوطة باللسانيات الحاسوبية، وهي التي تعمل على صياغة شكل الرسالة إذ تزود النظام بمختلف الضوابط والقيود التي تعطي لمحتوى الرسالة ومعطياتها المفهومية شكلا لسانيا سليم البناء والتركيب ومقبولا في اللغة المدروسة.

وبالتالي، سأحاول في هذا الفصل أن أنجز عملا تطبيقيا يمثل جزءا من المؤلفة التكتيكية للنظام، والتي ترجع مسؤولية إنشائها إلى اللسانيات الحاسوبية، هذه المؤلفة تمكن النظام من القواعد التركيبية والصرفية التي تضفي على محتوى الرسالة شكل الجمل العربية. وعليه، فإن هذا العمل التطبيقي يتضمن توليد جمل عربية بالاعتماد على البرمجة المباشرة؛ أي انطلاقا من معطيات وقواعد تركيبية وصرفية تُدخل إلى برنامج حاسوبي

فيقوم بعملية التوليد بطريقة آلية، من أجل تحقيق ذلك يمكن الاعتماد على عدة لغات برمجية أختار منها لغة برولوغ ***Prolog***.

1. لغة برولوغ وتحليل العربية

تجدر الإشارة إلى أن **عبد ذياب العجيلي** قد يكون من أوائل الباحثين العرب الذين حاولوا تحليل اللغة العربية باستخدام لغة برولوغ، وذلك في كتابه "الحاسوب واللغة العربية" –لكني لم أتمكن مع الأسف من الحصول على هذا المرجع–، وقد تناول فيه معالجة العربية في مستوياتها التحليلية والتركيبية، والنحوية والصرفية، والدلالية والأسلوبية، والهجائية، بالإضافة إلى تعرضه للترجمة الفورية بين العربية وغيرها من اللغات[1].

بالإضافة إلى ذلك، فإن هناك بعض الدراسات التي حاولت تشغيل اللغة العربية باستخدام لغة برولوغ، لكن بغرض صياغة المحللات التركيبية وليس بغرض صياغة المولدات الآلية للجمل. هذه الدراسة ترمي إلى إنشاء محلل للجمل الاسمية والفعلية العربية وسمي هذا المحلل بـــ**PANVS**، وهو يتألف من المكونات التالية:

- نظام تصافحي مع المستخدم؛
- محلل معجمي؛
- معجم؛

---

[1] ينظر: العناتي (وليد) والجبر (خالد)، دليل الباحث إلى اللسانيات الحاسوبية العربية، دار جرير، عمان، ط. 1، 2007، ص. 67.

- محلل صرفي؛
- مصحح للأخطاء الإملائية؛
- محلل تركيبي.

ومن أجل تمثيل مختلف المعارف النحوية العربية، تمت الاستعانة بصورانية **أنحاء العبارات المحددة *Definite-Clause Grammars***، والتي تسمح بكتابة الأنحاء المركبية في شكل عبارات منطقية بهدف تشغيل النظام باستخدام لغة برولوغ[1].

لقد اخترت لغة برولوغ من بين سائر لغات البرمجة المتوفرة في العمل التطبيقي نظرا لأن هذه اللغة تتمتع بعدة خصائص تجعلها مفضلة لدى الباحثين في معالجة الآلية للغات الطبيعية منها:

- تحتوي لغة برولوغ على هياكل بيانات[2] عريضة ومقعدة، يمكن توظيفها بسهولة في بناء البرامج وتغييرها، مما يجعل من السهل تمثيل البنى التركيبية والدلالية، والمداخل المعجمية؛
- يمكن للبرنامج المنشأ بهذه اللغة أن يختبر ويغير نفسه، تسمح هذه الخاصية باستعمال مناهج للبرمجة في غاية التجريد.

---

[1] ينظر: Al Daimi (J. Khalid) and Abdel-Amir (Maha A.), « *The syntactic analysis of arabic by machine* », Computers and Humanities, vol. 28, n° 1, Springer, Netherlands, 1994, p. 30.

[2] هيكل البيانات هو تركيب للبيانات تحدد به القيم التي تأخذها هذه البيانات وطريقة تخزينها والعمليات التي تتم عليها. معجم الحاسبات، مرجع سابق، ص. 57.

- برولوغ معدّ لتمثيل المعارف التي تمت صورنتها بالمنطق ذي الترتيب الأول، وامتدادات هذا المنطق هي سهلة التشغيل الآلي؛
- يستخدم برولوغ في كل أنواع المحللات دون استثناء؛
- تقنية التوحيد متضمنة في برولوغ، هذه التقنية التي يمكن أن تساهم في صياغة هياكل البيانات خطوة خطوة ولا يهم فيها ترتيب المراحل.

إن **Lisp** تشارك برولوغ في الميزتين الأوليين فقط، واللغات الأخرى مثل **Pascal** و**C** تفتقر إلى هذه الخصائص، لا يعني ذلك عدم إمكانية معالجة للغات الطبيعية باستخدام سائر لغات البرمجة، ولكن هناك لغة أنسب وأبسط من لغة للقيام بهذه العلمية[1].

1. النموذج المركبي للجملة العربية

تتألف الجملة من المركبات التي تتركب هي بدورها من وحدات لكل منها رتبة محددة. وقد بينت النظرية النحوية العربية بأن لكل وحدة لسانية موضعا داخل الجملة، وحتى لو تغير ترتيب تلك الوحدات أو موقعها داخل الجملة من أجل أهداف تداولية أو تواصلية أو نصية معينة، فإن مواضعها تبقى محفوظة.

إن الترتيب الأصلي للجمل يبقى مدار اختلاف بين اللسانيين؛ ففيما يعتبر **كرينبرك Greenberg** بأن الرتبة الأصلية في الإنجليزية هي 'فاعل فعل مفعول'، يقترح **مككولي McCawley** أن تكون الرتبة الأصلية فيها هي 'فعل فاعل مفعول'. وأما في العربية،

---

[1] ينظر: Covington (A. Michael), Natural language processing for Prolog programmers, Prentice-Hall, New Jersey, 2013, p. 12.

فيميل **كرينبرك Greenberg إلى** أن الرتبة الأصلية فيها هي 'فعل فاعل مفعول'، وتبعه على ذلك **الفاسي الفهري**، غير أن **تشومسكي Chomsky** ينكر وجود لغات من هذا النمط، اللهم إلا أن تكون محولة من الرتبة الأصلية 'فاعل فعل مفعول'[1].

وعلى ذلك يقوم الباحثون بتحليل الجملة العربية إلى شكلين: الجملة الاسمية التي تبدأ باسم، والجملة الفعلية التي تبدأ بفعل:

- جملة اسمية ← # اسم ..... #.

جملة فعلية ← # فعل ..... # [2].

يمكن القول بالنتيجة:

- إن الجملة الفعلية هي التي يعاد كتابتها إلى مركب فعلي ومركب اسمي:

جملة فعلية ← م فعلي م اسمي

- وإن الجملة الاسمية هي التي يعاد كتابتها إلى مركبين اسميين:

جملة اسمية ← م اسمي م اسمي

بالمقابل، يبدو لي أنه بالإمكان تحليل الجمل العربية بانتهاج منحى آخر أكثر تجريدا مما اقترحه الباحثون وذلك بالاعتماد على نموذج الحاج صالح الآنف الذكر:

[(ع ← م1) ± م2] ± خ

---

[1] ينظر: الفاسي الفهري (عبد القادر)، اللسانيات واللغة العربية، مرجع سابق، ص. 105.

[2] زكريا (ميشال)، الألسنية التوليدية والتحويلية وقواعد اللغة العربية، مرجع سابق، ص. 24.

إن هذا النموذج يشمل كل أنواع الجمل العربية مهما كان نوعها (الفعلية والاسمية والظرفية...) ومهما احتوت عليه من عناصر ووحدات؛

بحيث إن الجملة الاسمية هي التي تكون فيها العناصر التالية مثلا:

- ع = o ← السماء صافية؛
- ع= إن ← إن الصدقة تمحو الخطيئة.

والجملة الفعلية التي تتكون من عدة عناصر لكن يكون العامل فيها فعلا مثلا:

- ع = تكتب ← تكتب التلميذة على السبورة.
- ع= تذهب ← تذهب الفرحة الحزن...

وبالتالي، أقترح بأن يكون الترتيب الأصلي للجملة العربية على الشكل: 'عامل معمول1 معمول2 مخصص'، وأن أبسط تحققات هذا الترتيب تكون في الجملة الاسمية المكونة من مبتدأ وخبر، ومنها تتفرع كل التحققات الأخرى الممكنة، وذلك لأنه في هذه الجملة ينعدم العامل ويمكن الاستغناء عن المخصص، بل على المعمول2 أيضا، وتحتوي المركبات فيها على أقل ما يمكن من العناصر مثلا: في قوله تعالى: ﴿ وَلَئِن سَأَلْتَهُم مَنْ خَلَقَهُمْ لَيَقُولُنَّ اللهُ ﴾ (الزخرف، 87)، الشاهد في هذه الآية الكريمة هو الإتيان بجملة اسمية لا تحتوي إلا على مركب اسمي هو المبتدأ مكوَّن من اسم علم "الله"، وهذا في بعض التأويلات.

وعليه وبالاعتماد على الأنحاء الصورية المستخدمة ضمن المعالجة الآلية للغات الطبيعية في تحليل الجمل، يمكن اعتبار القواعد التي تولد الجمل العربية قواعد صورية يرمز لها بـــ **G**، وهي تنتج كل الجمل العربية السليمة من حيث التركيب أي جملا قواعدية ليس غير، وهي معرفة كالآتي:

$$G = (\{P, n, sg, g, g1, g2, s\}, \{\text{إن, سلوى, حلوة, كثيرا}\}, R, P)$$

حيث:

- $\{P, n, sg, g, g1, g2, s\}$[1]: هي مجموعة الرموز اللانهائية أو مجموعة المتغيرات.
- { إن ,سلوى ,حلوة ,جدا }: هي مجموعة الرموز النهائية.
- R: هو مجموعة منتهية من قواعد إعادة الكتابة.
- P: هو العنصر الابتدائي.

[1] إن هذه الرموز استخرجتها من المصطلحات التالية: **n:** noyau النواة، **sg:** syntagme gouvernant مركب عاملي، **g**: عامل،**g1**: معمول1، **g2**: معمول2، **s:** spicificateur مخصص.

وقواعد إعادة الكتابة هي على الشكل الموالي:

R = { P → n s;

n → sg g2;

sg → g g1;

g → إن ;

g1 → سلمى ;

g2 → حلوة ;

s → جدا }

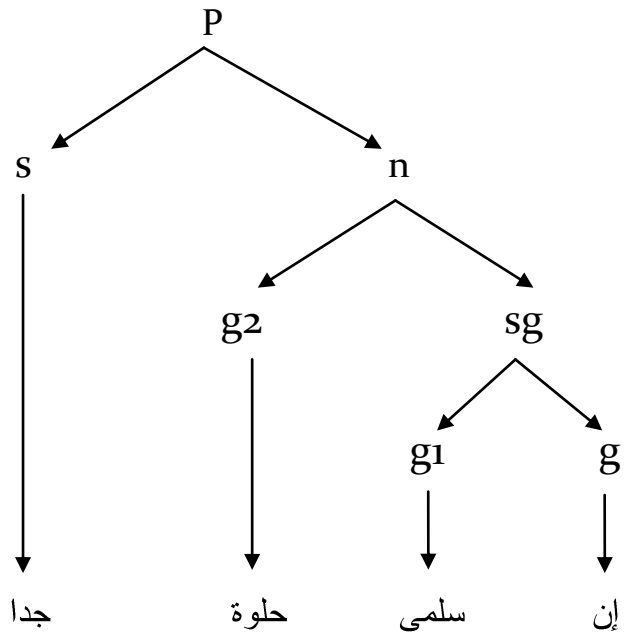
النحو **G** المولد للجمل العربية هو نحو صوري من الصنف 2 أو نحو لاسياقي حسب تراتبية تشومسكي؛ بما أن إنتاج اللغة[1] يتم بالتحديد عبر رمز غير نهائي عن اليسار وسلسلة غير معدومة عن اليمين، وأن هذه العملية تتم في كل السياقات وليس فقط في هذا السياق، فكل sg يمكن إعادة كتابته على شكل g g1 مثلا في أي سياق ورد.

إن الأنحاء التي من الصنف 2 تسمح بالتعرف على الجمل القواعدية (السليمة من حيث التركيب)، وبتمثيل هذه الجمل في مشجرات، كما أن لها تأويلا إثباتيا يتلاقى مع طبيعة لغة برولوغ البرمجية، والتي تتضمن إمكانية وصف بنية العناصر اللسانية بمعزل عن عملية المعالجة.

---

[1] اللغة هنا هي مجموعة السلاسل المنتهية من العناصر (الجمل القواعدية) *T، والتي تنتمي بدورها إلى المجموعة T التي تضم الرموز النهائية للغة. ينظر: Ligozat (Gérard), Représentation des connaissances et linguistique, *op.cit*, p. 36.

وبالتالي يمكن تمثيل الجملة السابقة بحسب المشجر التركيبي الآتي:



وتحتاج عملية التشغيل الآلي لتوليد الجمل إلى تحويل هذه البنى الجبرية المتمثلة في النموذج المركبي إلى بنى منطقية من أجل أن تسهل برمجتها بلغة برولوغ على النحو الموالي:

نحو العبارات المحددة DCG	**النحو المركبي SG**
gn(X):- art(Y), n(Z), concat(Y,Z,X).	GN → Art N

- في الخانة الأولى:

GN: مركب اسمي، Art: محدد، N: اسم، →: رمز قاعدة إعادة الكتابة، مثلا "الــبنت"؛

- في الخانة الثانية:

هذه العبارة المنطقية معناها أن سلسلة من الكلمات X هي مركب اسمي إذا وجد هناك سلسلة Y تكون محددا، وسلسلة Z تكون اسما، وإذا كانت X هي التي تصل بين Y وZ.

وقبل القيام بتحويل الجمل العربية إلى عبارات منطقية لتتم برمجتها بلغة برولوغ، أعرض تعريفا لهذه اللغة وأبرز عناصرها ووظائف تلك العناصر.

2. لغة برولوغ

لغة برولوغ صممت في بدايات سنوات 1970 على يد **كولمرووور Colmerauer** وفريقه في جامعة مرسيليا، كان الغرض من تصميمها هو برمجة نظام للاتصال بين الإنسان والآلة، وللمعالجة الآلية للغات الطبيعية. ترتكز هذه اللغة على المنطق الصوري خاصة على حساب المحمولات من الترتيب الأول، لذلك كانت تسميتها نابعة من التركيب بين المقطعين الأولين من البرمجة المنطقية أي "**Prolog**" من " PROgrammation LOGique"[1]. وقد كتبت لهذه اللغة منذ تصميمها عدة مترجمات ذات قواعد تركيبية مختلفة، من أشهرها **SWI-Prolog**[2] وهي التي أعتمد عليها عند إعداد برامج توليد الجمل في هذا الجانب التطبيقي.

---

[1] ينظر: Stehlé (Jean-Luc) et Hochard (Pierre), Ordinateurs et langages, *op.cit*, p. 117

[2] ينظر: www.swi-prolog.org

يمكن اعتبار برولوغ لغة للتعبير عن المعارف[1]، وهو لغة للبرمجة مكيفة مع الحساب الرمزي، وتحويل المواضيع والمحمولات والعلاقات. بالإضافة إلى أن هذه اللغة لا تحتوي على تعليمات[2]، أو تعيينات أو حلقات[3]، وإنما تتضمن عبارات تدل على وقائع، وقواعد، وأسئلة أو طلبات. يوظف **مترجم برولوغ *Interpréteur Prolog*** قاعدة المعارف (الوقائع، القواعد) للإجابة عن الأسئلة.

- الوقائع: pere(omar,reda).
- الأسئلة: ?- pere(X,Y).

X= omar.

Y= reda.

- القواعد: تسمح بتعريف العلاقة 'الجد' مثلا.

grand_pere(X,Z) :- pere(X,Y), pere(Y,Z).

مثلا :

pere(abdullah,mohamed).

pere(mohamed,fatima).

pere(mohamed,zayneb).

mere(fatima,hassane).

---

[1] يسمى هذا النوع من اللغات البرمجية باللغات الإثباتية حيث يبين فيها المشكل المراد حله، ولا تبين كيفية حله. وهناك أنواع أخرى كاللغات التي تحتوي على تعليمات (C CCCCC VV.....)، واللغات الوظيفية (Lisp, CMML...).

[2] **تعليمات *instructions*** بيان بعملية يطلب إلى الحاسوب تنفيذها، ويشمل نوع العملية وعناصرها أو براماتراتها. ينظر: معجم الحاسبات، مرجع سابق، ص. 101.

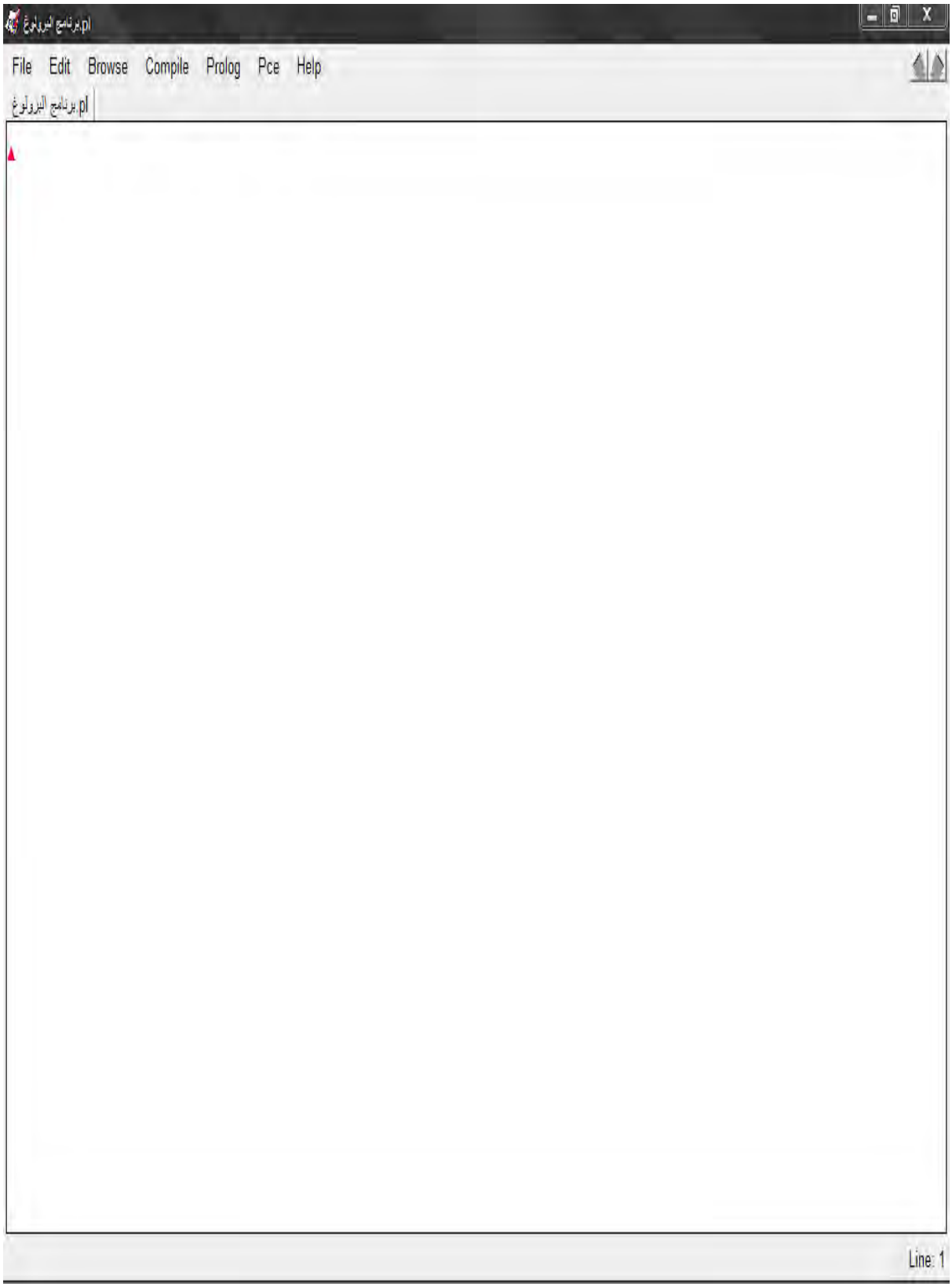
[3] **تعيين *affectation*** تخصيص قيمة لمتغير. **حلقة *boucle*** سلسلة من الأوامر لبرنامج تنفذ بطريقة تكرارية. ينظر: ccc iiaaaa eeffffff ormatique, *op.cit*, p. 4, p. 53.

```
mere(fatima,houssaine).
grand_pere(X,Z) :- pere(X,Y), mere(Y,Z).
grand_pere(X,Z) :- pere(X,Y), pere(Y,Z).
?- grand_pere(X,Y).
```

الإجابة:

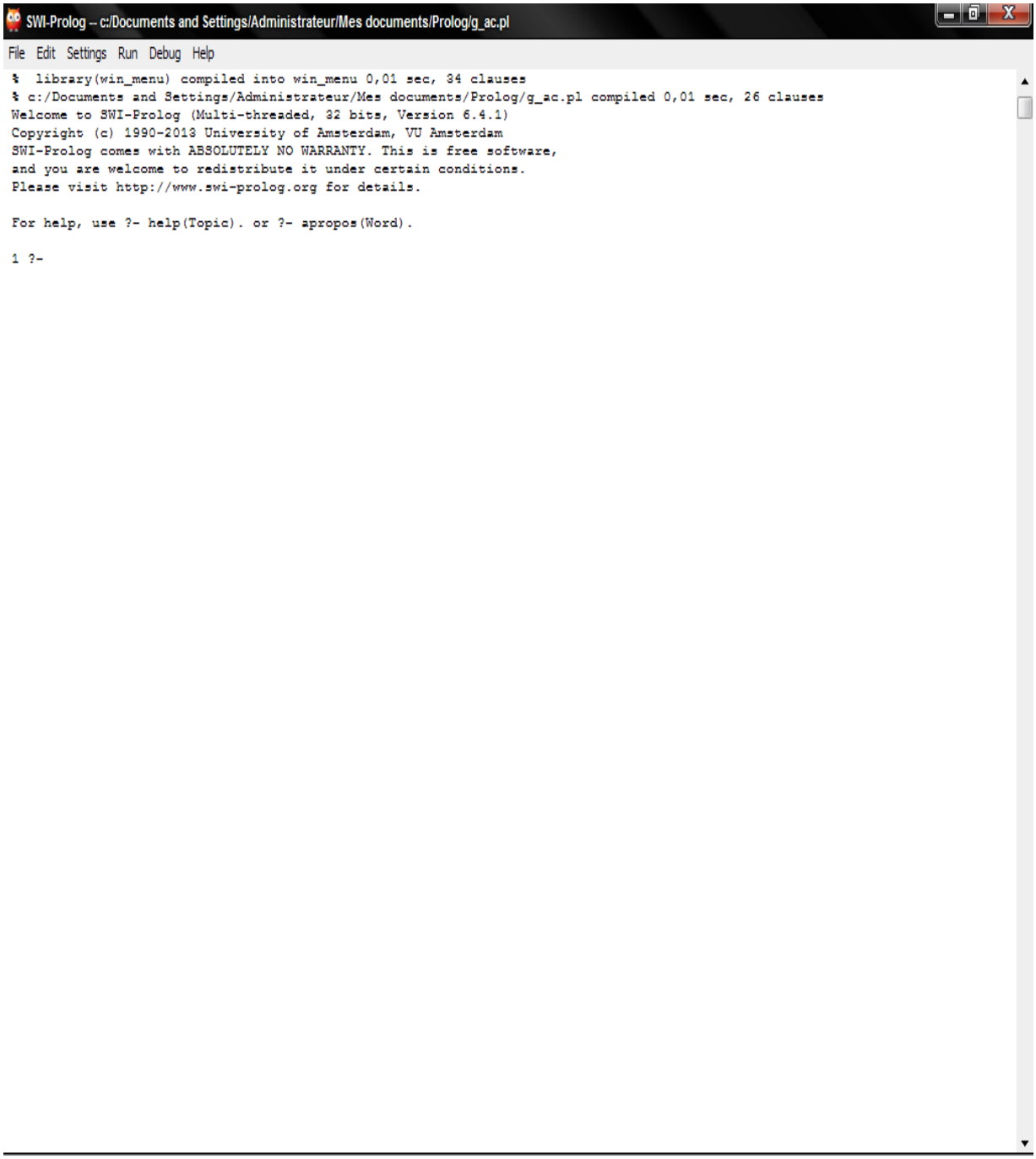
```
X = mohamed,
Y = hassane;
X = mohamed,
Y = houssaine;
X = abdullah,
Y = fatima;
X = abdullah,
Y = zayneb;
false.
```

الشكلان 15 و16 يمثلان على التوالي: المعالج النصي لبرنامج برولوغ، ومنفذ البرنامج.



الشكل15: ملف كتابة البرنامج ذو الامتداد *pl*

يمثل هذا الشكل الملف الذي تتم فيه كتابة الوقائع وإدخال قواعد المعطيات الخاصة بالبرنامج، وهو يتضمن التعليمة **compile** الذي يطلب من برنامج برولوغ ترجمة تلك الوقائع والقواعد المكتوبة بلغة منطقية إلى لغة الحاسوب السفلى حتى تتم معالجتها.



الشكل16: البرنامج المنفذ للبرولوغ

يمثل هذا الشكل برنامج برولوغ الذي يقوم بتنفيذ ما يطلب منه من أسئلة على حسب ما زود به من وقائع وقواعد، بعد أن تدخل إليه التعليمة **phrases -?**، فيعطي الإجابة المناسبة إذا لم تكن هناك أخطاء في كتابة البرنامج، وفي هذه الحالة يشير إلى وجود الأخطاء في السطر المناسب.

- **بعض عناصر اللغة**

- **الثابت *constant***: هو سلسلة من الحروف الخطية *caractères* المبدوءة بحروف أبجدية لاتينية صغيرة *lettres miniscules*، أو بأعداد. مثلا: a, algerie, 2014.
كذلك يمكن للبرنامج أن يتعرف على السلاسل المكتوبة بالحروف العربية ويعتبرها ثوابت.
- **المتغير *variable***: هو سلسلة من الحروف الخطية المبدوءة بحروف أبجدية لاتينية كبيرة *lettres majiscules*، أو بالرمز ' _ '. مثلا: X, A, _.
كذلك يمكن للبرنامج أن يتعرف على السلاسل المكتوبة بالحروف العربية المتغيرة بإضافة الرمز ' _ '، مثلا: إن = عامل_.
- **الحد *terme***: هو ثابت، أو متغير، أو رمز دالي مطبق على قائمة حدود. مثلا: X, a, f(a), a+3, 3+4.
- **الحرفي *littéral***: هو رمز المحمول (العلاقة) مطبق على قائمة حدود (يمكن أن تكون القائمة خالية)، وهو يعبر عن خاصية يمكن أن تكون صحيحة أو خاطئة.

مثلا: p(a), q(X), père(omar,Y), true, false. ينبغي عدم ترك فراغ بين المحمول والقوسين.

- عبارة برولوغ ***clause*** هي سلسلة مرتبة من الحرفيات المركبة، وقد تكون موجبة أو سالبة. مثلا:

pere(X,Y). عبارة موجبة

:- pere(X,Y), pere(Y,Z). عبارة سالبة

- برنامج برولوغ هو مجموعة من العبارات غير السالبة[1].

3. نحو العبارات المحددة للجملة العربية

بالانتقال إلى تمثيل الجمل العربية ذات الشكل: 'عامل معمول1 معمول2' باستخدام نحو العبارات المحددة **DCG**، أحاول تطبيق الإجراءات المتبعة في لغة برولوغ على بعض الأمثلة من الجمل العربية:

relation(إن,الجو,حار).

relation(يكرم,المؤمن,ضيفه).

relation(كان, البحر,هادئا).

- إدخال الحالة العامة:

relation(_عامل, معمول1_,معمول2_).

[1] ينظر: Pastre (Dominique), Cours de Prolog, 1990/2000, http://www.math-info.univ-paris5.fr/~pastre/prolog/cours.pdf، (تاريخ الاطلاع: 2033/08/15).

- ثم أطلب من البرنامج إعطاء قيم لتلك المتغيرات من خلال السؤال الآتي:

?- relation(_عامل, معمول1_,معمول2_).

فيكون رده كالآتي[1]:

_إن = عامل,

_الجو = معمول1,

_حار = معمول2;

_يكرم = عامل

_المؤمن = معمول1,

_ضيفه = معمول2;

_كان = عامل,

_البحر = معمول1,

_هادئا = معمول2.

true[2].

- ثم أطلب من البرنامج ربط تلك العناصر في جمل:

?- relation(_عامل,_معمول1,_معمول2), write(_عامل), write(' '), write(_معمول1),

write(' '), write(_معمول2), nl.

---

[1] العبارات في الأصل مكتوبة في البرنامج الذي اعتمدت عليه بهذا الشكل: معمول2 = هادئا_ وهذا مخالف لطريقة برولوغ في الكتابة؛ إذ من المفروض أن تكون مكتوبة من اليسار إلى اليمين على الشكل الذي وضعته في المتن، كما هي الحال عند استخدام الرموز اللاتينية، ويبدو أن ذلك راجع إلى الكتابة بالحروف العربية.

[2] هذا **الحرفي *littéral*** معناه أن البرنامج قد استوفى كل الحالات المطلوبة، فإذا ضغطنا على الزر " ; "، فإن البرنامج يجيب **بـــالحرفي *false***؛ أي ليست هناك حالات أخرى ممكنة.

فتكون الإجابة:

إن الجو حار

إن = عامل_,

الجو = معمول1_,

حار = معمول2_;

يكرم المؤمن ضيفه

يكرم = عامل_,

المؤمن = معمول1_,

ضيفه = معمول2_;

كان البحر هادئا

كان = عامل_,

البحر = معمول1_,

هادئا = معمول2_.

4. توليد الجمل

سأقوم فيما يلي وبالاعتماد على برنامج برولوغ بتوليد جمل انطلاقا من أمثلة أختارها عشوائيا، لكن أتحرى فيها البساطة والبعد عن تفصيلات القواعد النحوية العربية، وأركز على الجملة الفعلية المتكونة من فعل وفاعل، وعلى الجملة الاسمية المتكونة من مبتدأ وخبر (اسم مفرد أو جملة فعلية)[1].

[1] قد أثبت في الملحق كل البرامج التي تم تنفيذها في توليد مختلف الجمل كاملة.

في الحقيقة، هذا العمل التطبيقي الذي أقوم به لا يتناول كل عناصر النظام المولد، ولكنه يتناول جزءا منه وهو الذي يتعلق بتكوين الجمل وتأليفها انطلاقا من المعطيات التي تقدم إلى برنامج برولوغ، أما نظام التوليد فهو أكثر تعقيدا إذ ينطلق من التمثيلات الدلالية ليفرع نصوصا، وبالتالي فهو يحتوي على مؤلفتين استراتيجية وتكتيكية كما مر التفصيل فيهما.

1.4. توليد الجمل الاسمية

1- الجمل الاسمية المتكونة من مبتدأ وخبر يكون اسما مفردا: وهي أبسط صور التأليف في اللغة العربية إلى جانب الجملة الفعلية المتكونة من فعل وفاعل[1].

وحسب النموذج الخليلي: [(ع ← م1) ± م2] ± خ تكون كالآتي:

ع = Ø؛

م1: المبتدأ؛

م2: الخبر؛

خ = Ø.

أ- أقوم بداية بإدخال الوقائع وهي العناصر اللسانية التي ستكون الجملة الاسمية، فأرمز إلى المبتدأ بـــsn_1 ، وإلى الخبر بـــsn_2 .

[1] ينظر: الأنصاري (ابن هشام)، شرح قطر الندى وبل الصدى، ت. محمد محيي الدين عبد الحميد، مطبعة السعادة، مصر، ط. 11، 3963، ص. 44.

sn_1(محمد).

sn_1(سلمى).

sn_1(نور).

sn_1(الشمس).

sn_1(الموقف).

sn_2(كرمم).

sn_2(كريمة).

sn_2(هد...).

sn_2(تهدة..).

sn_2(مزدحم).

sn_2(مشرقة).

ب- ثم أقوم بإدخال البيانات الخاصة بنوع تلك العناصر اللسانية genre (مذكر masculin أو مؤنث feminin) حتى تتم المطابقة بين المبتدأ والخبر، كما يمكن أن أزوده ببيانات العدد nombre (مفرد singulier، مثنى duel، جمع pluriel).

genre(.... ,masculin).

genre(يم. ,masculin).

genre(..مجت,masculin).

genre(سلمى,feminin).

genre(وقف..ا,masculin).

genre(. و.,feminin).

genre(مزدحم,masculin).

genre(كريمة,feminin).

genre(لشمس.,feminin).

genre(دة. مج,feminin).

genre(رقة.. ,feminin).

ج- وبعد ذلك، أحدِّد القواعد الخاصة بالمطابقة بين المبتدأ والخبر في النوع:

```
accord(X,Y) :- genre(X,Z), genre(Y,Z).
```

كما هو معلوم: تكون المطابقة بين المبتدأ والخبر في النوع والعدد إلا في بعض الحالات.

د- أحدِّد قواعد توليد الجمل، وقواعد كتابتها، وتعميم هذه القواعد على كل الجمل الممكنة.

```
phrase(X,Y) :- sn_1(X), sn_2(Y), accord(X,Y).
phrase1(_) :-  phrase(X, Y), write(X), write('  '), write(Y), nl.
phrases :- findall(X, phrase1(X), _).
```

وأخيرا، أطلب من البرنامج تكوين الجمل.

```
?- phrases.
```

فيكون رد البرنامج كالآتي:

لآد كرملآ
لآلآلآ لآلآتهد
لآلآلآ لآلآم
لآى لآلآلآة
لآى لآلآلآدة
لآلآلآى لآلآلآ
نور لآلآيمة
نور مجتهدة
نور مشرقة
الشمس كيلآلآ
الشمس لآلآلآ

لآلشمس مشلآلآ
لآلآمولآلآ لآلآمم
لآلآمولآلآ لآلآآهلآ
لآلآموقف لآلآدحم

الملاحظ بأن البرنامج قد قام بتكوين كل الجمل الممكنة انطلاقا من البيانات التي قدمت إليه، وبالطريقة التي أرشد إليها من خلال القواعد، وبالتالي ولَّد جملا سليمة التركيب، لكن اللافت للانتباه أن بعض تلك الجمل سليمة التركيب غير مقبول لدى المتكلمين نظرا لعدم صحة الارتباط بين السمات الدلالية بين المبتدأ والخبر:

محمد مزدحم
الشمس لآلآمملآ
الشمس لآلآلآآهدة
المو لآ كرمم
لآلآمولآلآ لآلآتهد
لآلآمولآلآ لآلآدحم

من أجل تجاوز هذا المشكل أضيف إلى قائمة الوقائع بيانات تتعلق بالحركة أو الحياة (animé, non_animé) حتى يتفادى البرنامج تكوين تلك الجمل المتقدمة[1].

[1] ومع ذلك تظل نمذجة الجوانب الثقافية والاعتقادية في غاية الصعوبة حتى نتمكن من تجنب تكوين جمل من أمثال: الصادقون زائلون، الظالمون مأجورون...

trait(.محمد,masculin,animé).

trait(كرلي,masculin,animé).

trait(.مجتهد,masculin,animé).

trait(مم.. ,feminin,animé).

trait(.. الموز,masculin,singulier,non_animé).

trait(نور,feminin,animé).

trait(مزدحم,masculin,singulier,non_animé).

trait(يمة.. ,feminin,animé).

trait(الشمس,feminin,singulier,non_animé).

trait(..مجتهد,feminin,animé).

trait(مشرقة,feminin,singulier,animé).

trait(رقة.. ,feminin,singulier,non_animé).

- وهذه هي صورة إعداد البرنامج الخاص بهذا النوع من الجمل:

sn1sn2.pl

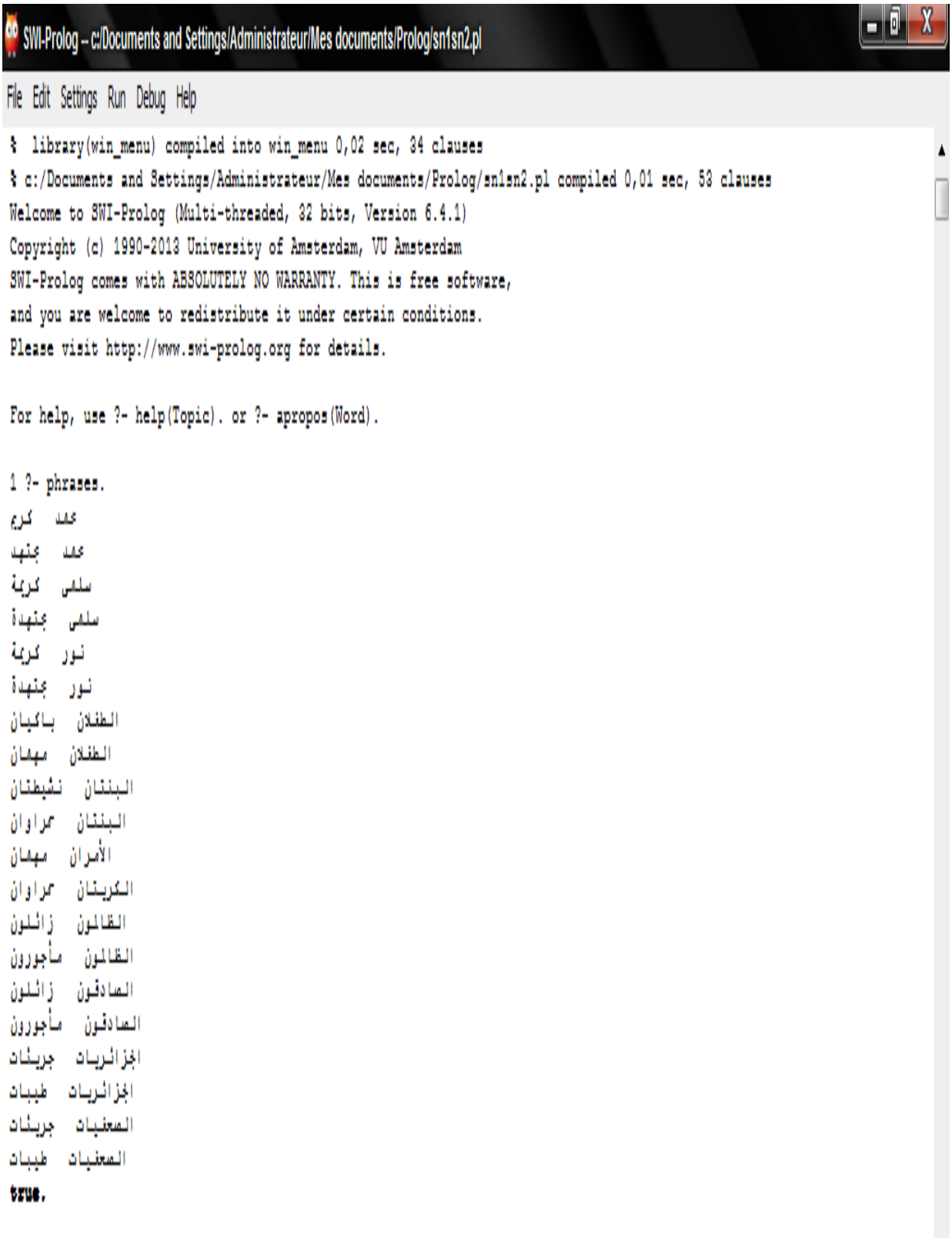
File Edit Browse Compile Prolog Pce Help

sn1sn2.pl

```
sn_1(محمد).
sn_1(سلمى).
sn_1(نور).
sn_1(الطفلان).
sn_1(البنتان).
sn_1(الأمران).
sn_1(الكريتان).
sn_1(الظالمون).
sn_1(الصادقون).
sn_1(الجزائريات).
sn_1(الصحفيات).
sn_2(كريم).
sn_2(كريمة).
sn_2(مجتهد).
sn_2(مجتهدة).
sn_2(باكيان).
sn_2(نشيطتان).
sn_2(مهمان).
sn_2(حمراوان).
sn_2(زائلون).
sn_2(مأجورون).
sn_2(جريئات).
sn_2(طيبات).
trait(محمد,masculin,singulier,animé).
trait(كريم,masculin,singulier,animé).
trait(مجتهد,masculin,singulier,animé).
trait(سلمى,féminin,singulier,animé).
trait(مجتهدة,féminin,singulier,animé).
trait(نور,féminin,singulier,animé).
trait(كريمة,féminin,singulier,animé).
trait(الطفلان,masculin,duel,animé).
trait(باكيان,masculin,duel,animé).
trait(الأمران,masculin,duel,non_animé).
trait(مهمان,masculin,duel,animé).
trait(مهمان,masculin,duel,non_animé).
trait(الظالمون,masculin,pluriel,animé).
trait(زائلون,masculin,pluriel,animé).
trait(الصادقون,masculin,pluriel,animé).
trait(مأجورون,masculin,pluriel,animé).
trait(البنتان,féminin,duel,animé).
trait(نشيطتان,féminin,duel,animé).
trait(الكريتان,féminin,duel,non_animé).
trait(حمراوان,féminin,duel,animé).
trait(حمراوان,féminin,duel,non_animé).
trait(الجزائريات,féminin,pluriel,animé).
trait(طيبات,féminin,pluriel,animé).
trait(الصحفيات,féminin,pluriel,animé).
trait(جريئات,féminin,pluriel,animé).
accord(X,Y) :- trait(X,W,Z,V), trait(Y,W,Z,V).
phrase(X,Y) :- sn_1(X), sn_2(Y), accord(X,Y).
phrase1(_) :-  phrase(X, Y), write(X), write('   '), write(Y), nl.
phrases :- findall(X, phrase1(X), _).
```

user:trait/4: (loaded) 25 clauses, number_of_rules(0) Line: 27

- وهذه هي صورة إجابة البرنامج:



2- **الجمل الاسمية المبدوءة بحرف ناسخ:** وهي الجمل الاسمية التي دخل عليها حرف ناسخ ينصب المبتدأ ويرفع الخبر، وهي: إنَّ، ليت، لكنّ، لعل، ولم أذكر "أنَّ" لأن لها أحكاما خاصة، بالإضافة إلى الجمل الاسمية المبدوءة بالحروف المشبهات بـــ"ليس"، إن، ما، لا التي ترفع المبتدأ وتنصب الخبر، لكن أقتصر على أمثلة من النوع الأول. وبالتالي، يكون النموذج على هذا الشكل:

ع = { إن، ليت، لكنّ، لعل}؛

م1: اسمها؛

م2: خبرها،

خ = Ø.

- أضيف إلى الوقائع المتقدمة البيانات الخاصة بالنواسخ، كما أدخل الأشكال الجديدة للأسماء السابقة: إن عامل نصب ← g_ac(إن)

g_ac(إن).

sn_1(الطفلان).

sn_1(الطفلين).

sn_2(باكيان).

sn_2(باكيين).

trait(.. ,masculin,singulier,animé,accusatif).

trait(إن,feminin,singulier,animé,accusatif).

trait(إن,masculin,duel,animé,accusatif).

```
trait(إن,feminin,duel,animé,accusatif).
trait(.. ,masculin,pluriel,animé,accusatif).
trait(ن.,feminin,pluriel,animé,accusatif).
trait(. ,masculin,singulier,non_animé,accusatif).
trait(إن,feminin,singulier,non_animé,accusatif).
trait(إن,masculin,duel,non_animé,accusatif).
trait(. ,feminin,duel,non_animé,accusatif).
trait(إن,masculin,pluriel,non_animé,accusatif).
trait(إن,feminin,pluriel,non_animé,accusatif).
trait(الطفلان,masculin,duel,animé,nominatif).
trait(الطفلين,masculin,duel,animé,accusatif).
trait(باكهان,masculin,duel,animé,nominatif).
trait(باكيين,masculin,duel,animé,accusatif).
```

- الملاحظ هنا هو إضافة بيانات جديدة تتعلق بالحالات الإعرابية (الرفع nominatif، النصب accusatif، الجر datif) نظرا إلى أن دخول الناسخ يغير آخر الاسمين.
- أقوم بإدخال القواعد الخاصة بالمطابقة، وتوليد الجمل، وتبقى قواعد الكتابة على حالها:

```
accord(X,Y,Z) :-
trait(X,A,B,C,accusatif),
trait(Y,A,B,C,accusatif),
trait(Z,A,B,C,nominatif).
phrase(X,Y,Z) :- g_ac(X), sn_1(Y), sn_2(Z), accord(X,Y,Z).
phrase1(_) :-  phrase(X, Y, Z), write(X), write('  '), write(Y), write('  '),
write(Z), nl.
```

phrases :- findall(X, phrase1(X), _).

- الإجابة:

?- phrases.

إن الطفلين باكمان

true.

- هذه صورة خاصة بإعداد البرنامج لهذه الحالة:

حرف ناصب.pl
File Edit Browse Compile Prolog Pce Help
حرف ناصب.pl
g_ac(إن).
g_ac(ليت).
g_ac(لكن).
g_ac(لعل).
sn_1(محمد).
sn_1(محمدا).
sn_1(سلمى).
sn_1(نور).
sn_1(نورا).
sn_1(الطفلان).
sn_1(الطفلين).
sn_1(البنتان).
sn_1(البنتين).
sn_1(الأمران).
sn_1(الأمرين).
sn_1(الكريتان).
sn_1(الكريتين).
sn_1(الظالمون).
sn_1(الظالمين).
sn_1(الصادقون).
sn_1(الصادقين).
sn_1(الجزائريات).
sn_1(الصحفيات).
sn_2(كريم).
sn_2(كريما).
sn_2(كريمة).
sn_2(مجتهد).
sn_2(مجتهدا).
sn_2(مجتهدة).
sn_2(باكيان).
sn_2(باكيين).
sn_2(نشيطتان).
sn_2(نشيطتين).
sn_2(مهمان).
sn_2(مهمين).
sn_2(حمراوان).
sn_2(حمراوين).
sn_2(زائلون).
sn_2(زائلين).
sn_2(مأجورون).
sn_2(مأجورين).
sn_2(جريئات).
sn_2(طيبات).

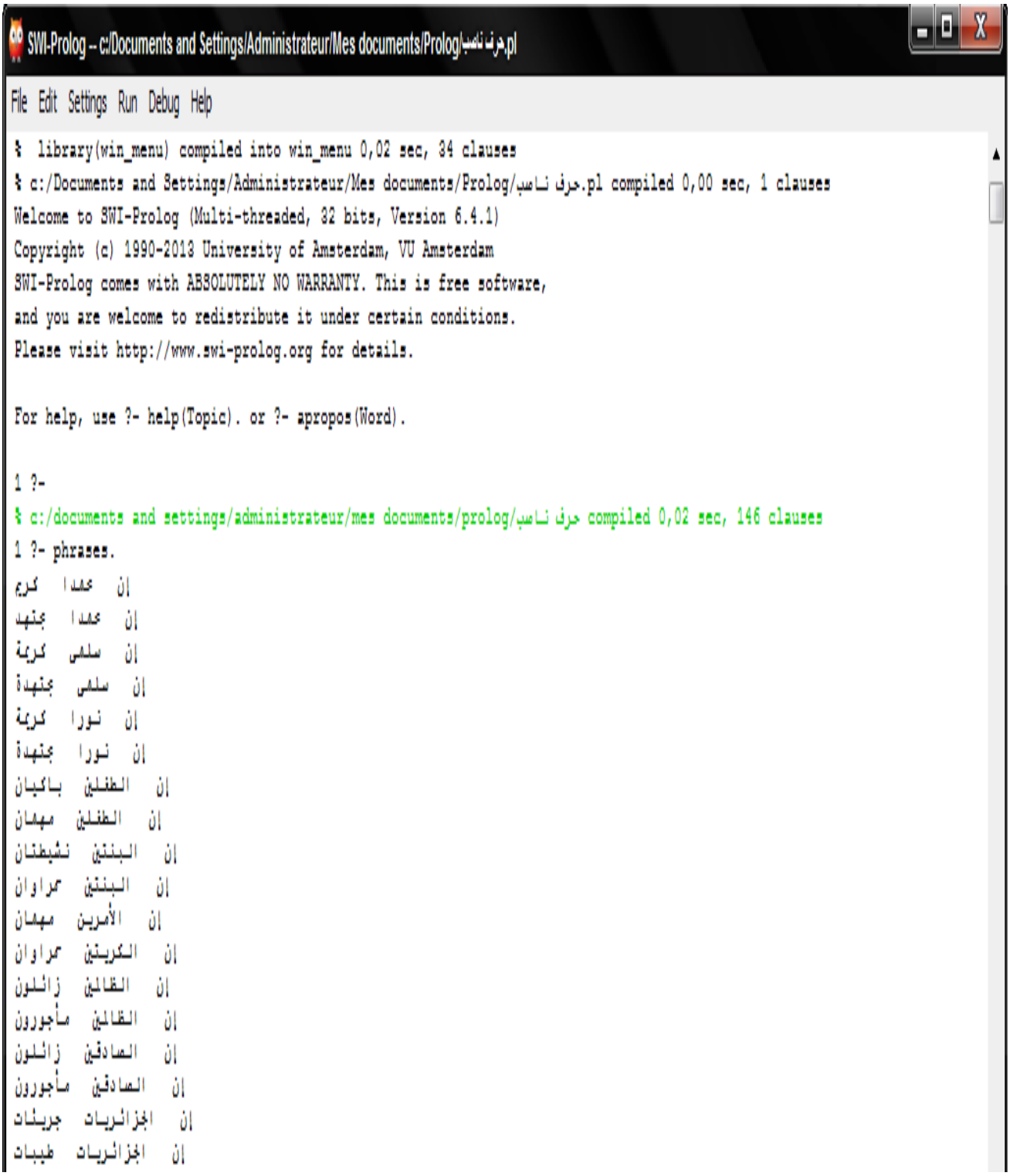
```
trait(إن,masculin,singulier,animé,accusatif).
trait(إن,feminin,singulier,animé,accusatif).
trait(إن,masculin,duel,animé,accusatif).
trait(إن,feminin,duel,animé,accusatif).
trait(إن,masculin,pluriel,animé,accusatif).
trait(إن,feminin,pluriel,animé,accusatif).
trait(إن,masculin,singulier,non_animé,accusatif).
trait(إن,feminin,singulier,non_animé,accusatif).
trait(إن,masculin,duel,non_animé,accusatif).
trait(إن,feminin,duel,non_animé,accusatif).
trait(إن,masculin,pluriel,non_animé,accusatif).
trait(إن,feminin,pluriel,non_animé,accusatif).
trait(لـيت,masculin,singulier,animé,accusatif).
trait(لـيت,feminin,singulier,animé,accusatif).
trait(لـيت,masculin,duel,animé,accusatif).
trait(لـيت,feminin,duel,animé,accusatif).
trait(لـيت,masculin,pluriel,animé,accusatif).
trait(لـيت,feminin,pluriel,animé,accusatif).
trait(لـيت,masculin,singulier,non_animé,accusatif).
trait(لـيت,feminin,singulier,non_animé,accusatif).
trait(لـيت,masculin,duel,non_animé,accusatif).
trait(لـيت,feminin,duel,non_animé,accusatif).
trait(لـيت,masculin,pluriel,non_animé,accusatif).
trait(لـيت,feminin,pluriel,non_animé,accusatif).
trait(لـعل,masculin,singulier,animé,accusatif).
trait(لـعل,feminin,singulier,animé,accusatif).
trait(لـعل,masculin,duel,animé,accusatif).
trait(لـعل,feminin,duel,animé,accusatif).
trait(لـعل,masculin,pluriel,animé,accusatif).
trait(لـعل,feminin,pluriel,animé,accusatif).
trait(لـعل,masculin,singulier,non_animé,accusatif).
trait(لـعل,feminin,singulier,non_animé,accusatif).
trait(لـعل,masculin,duel,non_animé,accusatif).
```

```
trait(لعل,feminin,duel,non_animé,accusatif).
trait(لعل,masculin,pluriel,non_animé,accusatif).
trait(لعل,feminin,pluriel,non_animé,accusatif).
trait(لكن,masculin,singulier,animé,accusatif).
trait(لكن,feminin,singulier,animé,accusatif).
trait(لكن,masculin,duel,animé,accusatif).
trait(لكن,feminin,duel,animé,accusatif).
trait(لكن,masculin,pluriel,animé,accusatif).
trait(لكن,feminin,pluriel,animé,accusatif).
trait(لكن,masculin,singulier,non_animé,accusatif).
trait(لكن,feminin,singulier,non_animé,accusatif).
trait(لكن,masculin,duel,non_animé,accusatif).
trait(لكن,feminin,duel,non_animé,accusatif).
trait(لكن,masculin,pluriel,non_animé,accusatif).
trait(لكن,feminin,pluriel,non_animé,accusatif).
trait(محمد,masculin,singulier,animé,nominatif).
trait(محمدا,masculin,singulier,animé,accusatif).
trait(كريم,masculin,singulier,animé,nominatif).
trait(كريما,masculin,singulier,animé,accusatif).
trait(مجتهد,masculin,singulier,animé,nominatif).
trait(مجتهدا,masculin,singulier,animé,accusatif).
trait(سلمى,féminin,singulier,animé,nominatif).
trait(سلمى,féminin,singulier,animé,accusatif).
trait(مجتهدة,féminin,singulier,animé,nominatif).
trait(مجتهدة,féminin,singulier,animé,accusatif).
trait(نور,féminin,singulier,animé,nominatif).
trait(نورا,féminin,singulier,animé,accusatif).
trait(كريمة,féminin,singulier,animé,nominatif).
trait(كريمة,féminin,singulier,animé,accusatif).
trait(الطفلان,masculin,duel,animé,nominatif).
trait(الطفلين,masculin,duel,animé,accusatif).
trait(باكيان,masculin,duel,animé,nominatif).
trait(باكيين,masculin,duel,animé,accusatif).
trait(الأمران,masculin,duel,non_animé,nominatif).
trait(الأمرين,masculin,duel,non_animé,accusatif).
trait(مهمان,masculin,duel,animé,nominatif).
```

```
trait(مهمين,masculin,duel,animé,accusatif).
trait(مهمان,masculin,duel,non_animé,nominatif).
trait(مهمين,masculin,duel,non_animé,accusatif).
trait(الظالمون,masculin,pluriel,animé,nominatif).
trait(الظالمين,masculin,pluriel,animé,accusatif).
trait(زائلون,masculin,pluriel,animé,nominatif).
trait(زائلين,masculin,pluriel,animé,accusatif).
trait(الصادقون,masculin,pluriel,animé,nominatif).
trait(الصادقين,masculin,pluriel,animé,accusatif).
trait(مأجورون,masculin,pluriel,animé,nominatif).
trait(مأجورين,masculin,pluriel,animé,accusatif).
trait(البنتان,féminin,duel,animé,nominatif).
trait(البنتين,féminin,duel,animé,accusatif).
trait(نشيطتان,féminin,duel,animé,nominatif).
trait(نشيطتين,féminin,duel,animé,accusatif).
trait(الكريتان,féminin,duel,non_animé,nominatif).
trait(الكريتين,féminin,duel,non_animé,accusatif).
trait(حمراوان,féminin,duel,animé,nominatif).
trait(حمراوين,féminin,duel,animé,accusatif).
trait(حمراوان,féminin,duel,non_animé,nominatif).
trait(حمراوان,féminin,duel,non_animé,accusatif).
trait(الجزائريات,féminin,pluriel,animé,nominatif).
trait(الجزائريات,féminin,pluriel,animé,accusatif).
trait(طيبات,féminin,pluriel,animé,nominatif).
trait(طيبات,féminin,pluriel,animé,accusatif).
trait(صحفيات,féminin,pluriel,animé,nominatif).
trait(صحفيات,féminin,pluriel,animé,accusatif).
trait(جريئات,féminin,pluriel,animé,nominatif).
trait(جريئات,féminin,pluriel,animé,accusatif).
accord(X,Y,Z) :-
trait(X,A,B,C,accusatif),
trait(Y,A,B,C,accusatif),
trait(Z,A,B,C,nominatif).
phrase(X,Y,Z) :- g_ac(X), sn_1(Y), sn_2(Z), accord(X,Y,Z).
phrase1(_) :-  phrase(X, Y, Z), write(X), write('   '), write(Y), write('   '), write(Z), nl.
phrases :- findall(X, phrase1(X), _).
```

Line: 116

*ثم هذه صورة إجابة البرنامج:


SWI-Prolog -- c:/Documents and Settings/Administrateur/Mes documents/Prolog/حرف ناصب.pl
File Edit Settings Run Debug Help
% library(win_menu) compiled into win_menu 0,02 sec, 34 clauses
% c:/Documents and Settings/Administrateur/Mes documents/Prolog/حرف ناصب.pl compiled 0,00 sec, 1 clauses
Welcome to SWI-Prolog (Multi-threaded, 32 bits, Version 6.4.1)
Copyright (c) 1990-2013 University of Amsterdam, VU Amsterdam
SWI-Prolog comes with ABSOLUTELY NO WARRANTY. This is free software,
and you are welcome to redistribute it under certain conditions.
Please visit http://www.swi-prolog.org for details.
For help, use ?- help(Topic). or ?- apropos(Word).
1 ?-
% c:/documents and settings/administrateur/mes documents/prolog/حرف ناصب compiled 0,02 sec, 146 clauses
1 ?- phrases.
إن محمدا كريم
إن محمدا مجتهد
إن سلمى كريمة
إن سلمى مجتهدة
إن نورا كريمة
إن نورا مجتهدة
إن الطفلين باكيان
إن الطفلين مهمان
إن البنتين نشيطتان
إن البنتين مراوان
إن الأمرين مهمان
إن الكريمتين مراوان
إن القالين زائلون
إن القالين مأجورون
إن الصادقين زائلون
إن الصادقين مأجورون
إن الجزائريات جريئات
إن الجزائريات طيبات

ليد محمدا كريم
ليد محمدا مجتهد
ليد سلمى كريمة
ليد سلمى مجتهدة
ليد نورا كريمة
ليد نورا مجتهدة
ليد الطفلين باكيان
ليد الطفلين مهمان
ليد البنتين نشيطتان
ليد البنتين مراوان
ليد الأمرين مهمان
ليد الكريتين مراوان
ليد القالين زائلون
ليد القالين مأجورون
ليد الصادقين زائلون
ليد الصادقين مأجورون
ليد الجزائريات جريئات
ليد الجزائريات طيبات
لكن محمدا كريم
لكن محمدا مجتهد
لكن سلمى كريمة
لكن سلمى مجتهدة
لكن نورا كريمة
لكن نورا مجتهدة
لكن الطفلين باكيان
لكن الطفلين مهمان
لكن البنتين نشيطتان
لكن البنتين مراوان
لكن الأمرين مهمان
لكن الكريتين مراوان
لكن القالين زائلون
لكن القالين مأجورون
لكن الصادقين زائلون
لكن الصادقين مأجورون
لكن الجزائريات جريئات
لكن الجزائريات طيبات
لعل محمدا كريم
لعل محمدا مجتهد
لعل سلمى كريمة
لعل سلمى مجتهدة
لعل نورا كريمة
لعل نورا مجتهدة
لعل الطفلين باكيان
لعل الطفلين مهمان
لعل البنتين نشيطتان
لعل البنتين مراوان
لعل الأمرين مهمان
لعل الكريتين مراوان
لعل القالين زائلون
لعل القالين مأجورون
لعل الصادقين زائلون
لعل الصادقين مأجورون
لعل الجزائريات جريئات
لعل الجزائريات طيبات
true.

3- **الجمل الاسمية التي خبرها جملة فعلية:** الجملة الاسمية التي يكون خبرها جملة هي التي يسميها النحاة بالجملة الكبرى، ويقابلها الجملة الصغرى وهي التي يكون خبرها مفردا[1]، والجملة المخبر بها قد تكون فعلية أو اسمية، والجملة الفعلية المخبر بها تشبه إلى حد كبير الجملة الفعلية مما جعل بعض النحاة يعدونهما شيئا واحدا[2]، إلا أن حالات المطابقة تبين أنهما مختلفان، وسأتعرض إليها فيما يلي.

أ- إدخال الوقائع الخاصة بالمكونات التي تكون فعلا sv أو مبتدأ sn.

sn(الولاد).

sn(سلمى).

sn(نور).

sv(يكبر).

sv(الولد).

sv(يجتهد).

sv(تجلد).

sn(الطفلان).

sn(الولدان).

sv(تسلعان).

sv(يصلان).

sn(المدرسون).

sv(ينشغلون).

sn(التلميذات).

---

[1] ينظر: الأنصاري (ابن هشام)، مغني اللبيب عن كتب الأعاريب، مرجع سابق، ج. 2، ص. 477.

[2] خاصة الكوفيين الذين يجوزون تقديم الفاعل على فعله، وبذلك تتماثل الجملة الفعلية والجملة الاسمية التي خبرها جملة فعلية عندهم. ينظر: المخزومي (مهدي)، مدرسة الكوفة ومنهجها في دراسة اللغة والنحو، مطبعة مصطفى الحلبي، مصر، ط.2، 1958، ص. 278. وتبعهم على ذلك بعض اللسانيين المحدثين.

sv(ينتبهن).

ب- توضيح بيانات المقولات المميزة trait (النوع: مذكر أو مؤنث، العدد: مفرد، مثنى، جمع) التي تتطابق فيها المكونات.

trait(محمدا,masculin,singulier,animé).
trait(ملاعم,masculin,singulier,animé).
trait(ملاتهد,masculin,singulier,animé).
trait(سلمى,feminin,singulier,animé).
trait(نولا,feminin,singulier,animé).
trait(تكرم,feminin,singulier,animé).
trait(تجتهد,feminin,singulier,animé).
trait(الطفلان,masculin,duel,animé).
trait(الطفلان,masculin,duel,animé).
trait(يصلان,masculin,duel,animé).
trait(مصلان,masculin,duel,non_animé).
trait(الا لاتان,feminin,duel,non_animé).
trait(الالالان,feminin,duel,animé).
trait(الالالا لا,feminin,duel,non_animé).
trait(اللالالون,masculin,pluriel,animé).
trait(إيتالالون,masculin, pluriel,animé).
trait(التلملات,feminin,pluriel,animé).
trait(منلاهن, feminin,pluriel,animé).

ج- إدخال القواعد الخاصة بالمطابقة، وتوليد الجمل، وقواعد كتابتها، وتعميم هذه القواعد على كل الجمل الممكنة، وأخيرا وضع السؤال.

```
accord(X,Y,Z) :- trait(X,A,B,C), trait(Y, A,B,C), trait(Z, A,B,C).
phrase(X,Y) :- sn(X), sv(Y), accord(X,Y,Z).
phrase1(_) :- phrase(X, Y), write(X), write('  '), write(Y), nl.
phrases :- findall(X, phrase1(X), _).
?- phrases.
```

د- الإجابة.

```
الالالاد يلارم
محمد  يلجلالا
سلمى  تكرلا
لالمى  تجتهد
نور  تكرلا
نور  تلاتهد
الطفلان  ملالا ل
لالارالجلان  تسرعان
المدرسد ول  ينشغلون
التلميذات  ينلاهم ل
true.
```

الملاحظ هنا أن البرنامج لم يولِّد إلا الجمل التي فيها تطابق بين المبتدأ والفعل من خلال الضمير العائد على المبتدأ، وهذا التطابق يكون في النوع والعدد، كما أنه لم يولِّد جملا غير مقبولة نظرا لتزويده بالبيانات الخاصة بالحركة.

2.4. توليد الجمل الفعلية

أتطرق هنا إلى الجملة الفعلية في أبسط صورها وهي المتكونة من فعل وفاعل.

وحسب النموذج الخليلي: [(ع ← م1) ± م2] ± خ تكون كالآتي:

ع = فعل (ليس بالضرورة أن يكون لازما)؛

م1: فاعل؛

م2 = Ø؛

خ = Ø.

أ- إدخال الوقائع الخاصة بالمكونات التي تكون فعلا sv أو فاعلا sn.

sn(محمد).

sn(سلمى).

sn(نور).

sv(يلارم).

sv(تكرلا).

sv(مجلهد).

sv(تلالهلا).

sn(الطفلان).

sn(الدراجلان).

sv(الالالاع.

sv(يصلا).

sn(المدرسون).

sv(ينشغل).

sn(التلميذات).

sv(تلابه).

ب- توضيح البيانات التي يتطابق فيها الفعل والفاعل.

genre(محمد,masculin,animé).

genre(علي,masculin,animé).

genre(عبدالله,masculin,animé).

genre(سلمى,feminin,animé).

genre(نورا,feminin,animé).

genre(نور,feminin,non_animé).

genre(تكرم,feminin,animé).

genre(تجتهد,feminin,animé).

genre(الطفلان,masculin,animé).

genre(يصل,masculin,animé).

genre(يصلا,masculin,non_animé).

genre(الرجلان,feminin,non_animé).

genre(تسرا, feminin,animé).

genre(تسرع, feminin,non_animé).

genre(المسلمون,masculin,animé).

genre(ينشدون,masculin,animé).

genre(المسلمات,feminin,animé).

genre(ينتبهن, feminin,animé).

ج- إدخال القواعد الخاصة بالمطابقة، وتوليد الجمل، وقواعد كتابتها، وتعميم هذه القواعد على كل الجمل الممكنة، وأخيرا وضع السؤال.

```
accord(X,Y) :- genre(X,Z), genre(Y,Z).
phrase(X,Y) :- sv(X), sn(Y), accord(X,Y).
phrase1(_) :- phrase(X, Y), write(X), write('  '), write(Y), nl.
phrases :- findall(X, phrase1(X), _).
?- phrases.
```

د- الإجابة.

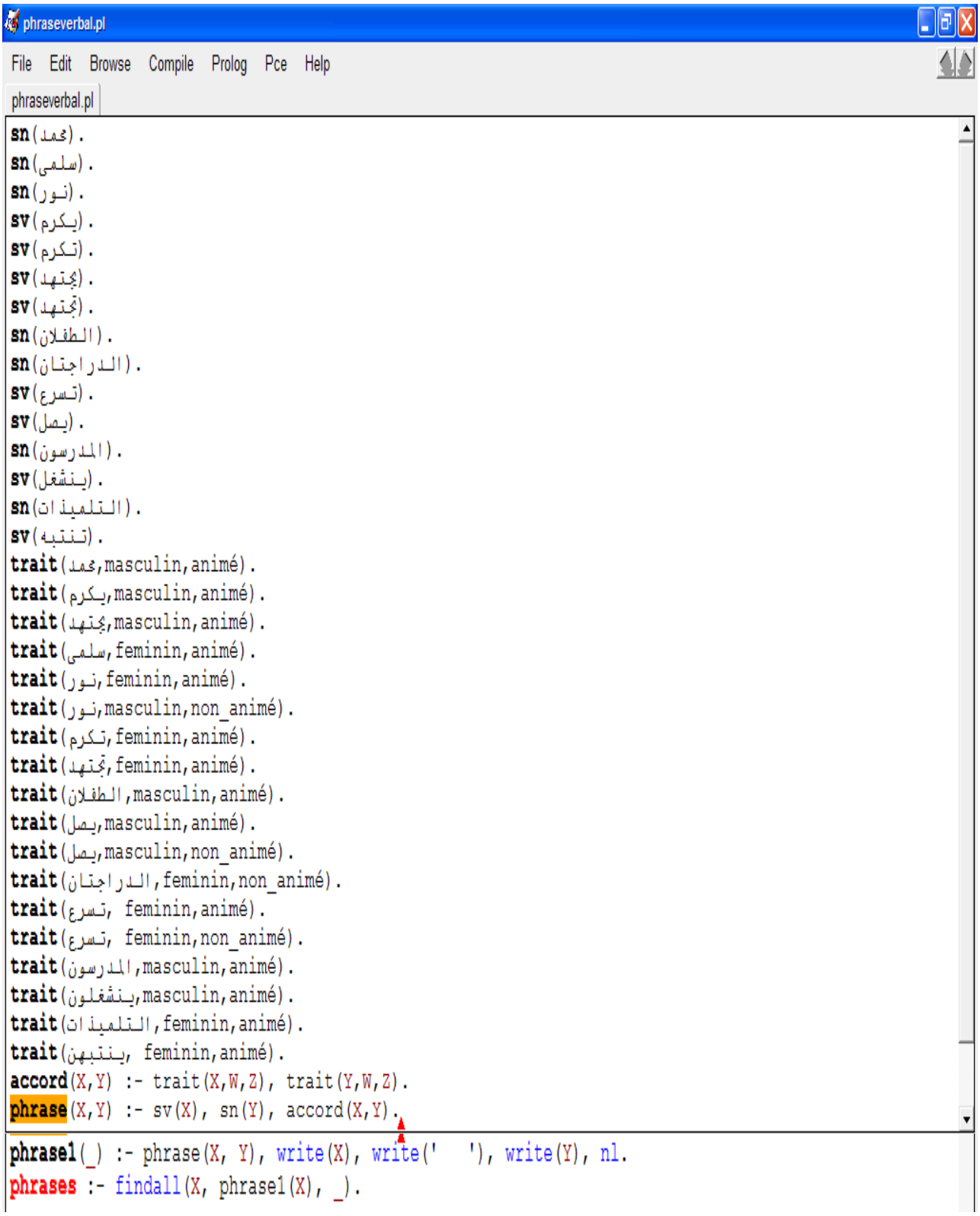
ملالام محمد
ملاللا الطفلان
ملالا المدرسون
تكرم سلمى
تكرم نور
لاللاقلاللاراجتان
لاللاللاملاللاقيذا لا
يجتهد محمد
يجلاهلا الطفلان
ملاتهد المدرسون
تجتهد سلمم
لاللاهلا نور
لاللاهلاللادلاللالان
لاتهلاللاللاملات
تسرع لاللامى
تسرع نولا
تسرعلاللادراجلان
تسرلا التلميلا لا
ملا محمد
ملالا الطفلان
يصل اللاللاون

إذن، قام البرنامج بتكوين كل الجمل الفعلية الممكنة، وطابق بين الفعل وفاعله في النوع والعدد، وذلك انطلاقا من القواعد التي زود بها، بيد أنه شكّل جملا غير مقبولة أيضا، وذلك لعدم تزويده بقيود دلالية تمنع تكوين مثل هذه الجمل.

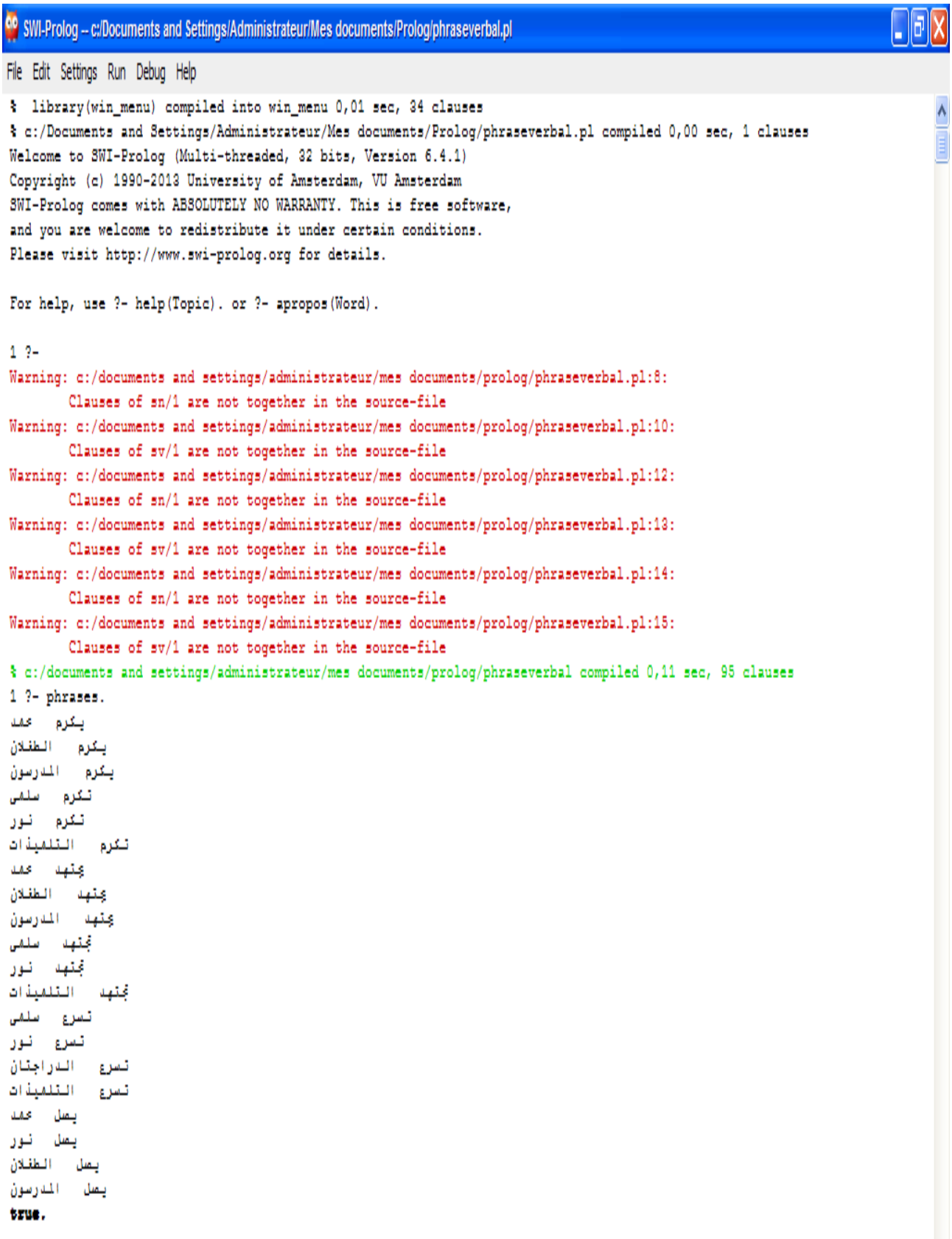
وتجدر الإشارة إلى أن هناك تشابها بين الجملة الفعلية والجملة الاسمية التي خبرها جملة فعلية، والفرق يظهر في أوجه المطابقة بحيث إنه في الجملة الفعلية لا تطابق إلا في النوع (التذكير والتأنيث)[1]، أما في الجملة الاسمية التي خبرها جملة فعلية فتكون المطابقة في النوع والعدد (مفرد، مثنى، جمع) بالإضافة إلى سمة الحركة طبعا.

- هذه صورة البرنامج الخاصة بالجمل الفعلية:

[1] على الغالب والمشهور من كلام العرب، ولا آخذ في الاعتبار لغة "أكلوني البراغيث"، التي يتطابق فيها الفعل والفاعل حتى في العدد.


phraseverbal.pl
File Edit Browse Compile Prolog Pce Help
phraseverbal.pl
sn(محمد).
sn(سلمى).
sn(نور).
sv(يكرم).
sv(تكرم).
sv(يجتهد).
sv(تجتهد).
sn(الطفلان).
sn(الدراجتان).
sv(تسرع).
sv(يصل).
sn(المدرسون).
sv(ينشغل).
sn(التلميذات).
sv(تنتبه).
trait(محمد,masculin,animé).
trait(يكرم,masculin,animé).
trait(يجتهد,masculin,animé).
trait(سلمى,feminin,animé).
trait(نور,feminin,animé).
trait(نور,masculin,non_animé).
trait(تكرم,feminin,animé).
trait(تجتهد,feminin,animé).
trait(الطفلان,masculin,animé).
trait(يصل,masculin,animé).
trait(يصل,masculin,non_animé).
trait(الدراجتان,feminin,non_animé).
trait(تسرع, feminin,animé).
trait(تسرع, feminin,non_animé).
trait(المدرسون,masculin,animé).
trait(ينشغلون,masculin,animé).
trait(التلميذات,feminin,animé).
trait(ينتبهن, feminin,animé).
accord(X,Y) :- trait(X,W,Z), trait(Y,W,Z).
phrase(X,Y) :- sv(X), sn(Y), accord(X,Y).
phrase1(_) :- phrase(X, Y), write(X), write(' '), write(Y), nl.
phrases :- findall(X, phrase1(X), _).

- ثم هذه صورة لإجابة البرنامج:



في نهاية هذا الفصل التطبيقي الذي حاولت من خلاله استخدام لغة برولوغ في توليد مجموعة من الجمل العربية، وهي على التوالي: الجملة الاسمية سواء التي يكون العامل فيها الابتداء أو حرفا ناسخا، والجملة الاسمية التي خبرها جملة فعلية، وأخيرا الجملة الفعلية.

وقد تضمنت تلك البرامج المنفذة في توليد الجمل تزويد الحاسوب بالبيانات اللازمة التي تمثل العناصر المكونة للجمل، وكذا البيانات التي تمكن من المطابقة بين تلك العناصر من أجل تشكيل جمل سليمة البناء والتركيب، وقد التزمت في ذلك بالبساطة والاختصار إذ لم أتبع إلا الترتيب الأصلي للعناصر الأساسية في الجمل، ولم أذكر الجمل الموسعة التي تتكون من مخصصات، وكذا مختلف الأحوال التي تكون عليها تلك العناصر.

خـــــاتمــة

خـــاتـــــمـــة

في ثنايا الفصول الخمسة التي تضمنتها هذه الرسالة، حاولت أن أوضح مختلف المفاهيم المتصلة بالموضوع مثل مفاهيم الجملة واللسان والتوليد، كما أنني تعرضت إلى بيان عناصر الدراسة اللسانية الحاسوبية وبيان حدودها ومنهجها وأهدافها من دراسة الظاهرة اللغوية، وحاولت الاستفادة من تلك الدراسة في شرح عملية توليد الجمل التي أنشئت حولها الكثير من النظريات المتعددة المشارب تتراوح بين الصورية والوظيفية والمعرفية. ثم تطرقت إلى أسس التوليد الآلي للجمل في الفصل الرابع، وختمت دراستي بالجانب التطبيقي المتضمن صياغة برامج حاسوبية لتوليد الجمل في اللسان العربي باستخدام لغة برولوغ، بعد أن تناولت بالعرض زمرة من المولدات الآلية التي تعتبر سندا ومعتمدا في هذه العملية.

وقد قادتني هذه الدراسة لهذا الموضوع المتداخل التخصصات إلى مجموعة من النتائج أذكرها فيما يلي:

1. إن مفهومي الجملة والكلام يصعب تحديدهما بشكل قاطع، لكن كثيرا من النحاة والدارسين المحدثين قد تفطنوا إلى وضع حدود إجرائية بين تلك المفاهيم، كربطهم مفهوم الجملة بشكل القول وهيكله، وربطهم مفهوم الكلام بالقول موظفا في عملية الاتصال.

2. يمكن أن يكون نموذج الحاج صالح للجملة العربية منطلقا للدراسات اللسانية الحاسوبية العربية لعدة أسباب من أهمها أنه يتميز بصوريته في تمثيل الجمل العربية، وشموليته لجميع أنواعها والعناصر التي يمكن أن تضمها، واعتماده على فكرة العامل الذي يتلاقى فيها مع أنحاء التعلق.

3. أقترح، تبعا للحاج صالح، أن يكون الترتيب الأصلي للجملة العربية على الشكل: 'عامل معمول1 معمول2 مخصص'، وأن يستخدم هذا الترتيب في تحليل الجملة العربية ومعالجتها آليا.

4. إن الأدبيات العربية التي تناولت اللسانيات الحاسوبية تعتمد على ربط اللسانيات أو علوم اللغة بعلوم الحاسوب، على الرغم من أن هناك مجالات كثيرة دقيقة وصعبة الفصل بينها ضمن اللسانيات الحاسوبية، إلا أن الخلط بين المصطلحات والمفاهيم لا يزال طاغيا على تلك الأدبيات مثل: الهندسة اللسانية، تكنولوجيا اللسان... لذلك المطلوب من الباحثين المصطلحيين والمعربين إلى بذل المزيد من الجهد في كشف اللبس بينها.

5. يبدو أن مصطلح "اللسانيات الحاسوبية" المستخدم في الأدبيات العربية هو مصطلح غامض لأنه يشمل عدة محاور تلتقي فيها الدراستان اللسانية والحاسوبية، مما يصعب على القارئ العربي التفريق بين تلك المحاور المتعددة داخل ذلك التماس.

6. إن الجهاز المصطلحي الفرنسي يشتمل على مصطلحين للدلالة على تداخل الدراستين اللسانية والحاسوبية وهما ***linguistique informatique*** و***linguistique computationnelle*** لذلك يمكن في نظري أن يترجما إلى مصطلحين مختلفين هما "**اللسانيات المعلومية**" و"**اللسانيات الحاسوبية**" على التوالي خاصة أن المصطلحين ليسا بمترادفين عند بعض الباحثين الفرنسيين.
7. هناك عدة نقائص قد لاحظها الباحثون على أنظمة التوليد المتوفرة؛ فعادة ما تكون تلك الأنظمة مرتبطة بتطبيق خاص. كما أنها غالبا ما تكون متصلة بلغة معينة قد يصعب تعميمها على لغات أخرى، وتتميز النصوص التي تنتجها بشكلها الجامد. بالإضافة إلى ذلك، لا يرتكز الجانب اللساني لأنظمة المدروسة على نموذج لساني إلا جزئيا.
8. إن العمل التطبيقي الذي أنجزته تناول موضوعا لسانيا هو توليد الجمل العربية وقام ببرمجة هذه العملية بالاعتماد على لغة برولوغ، لذا فإن هذا البحث ذو صبغة لسانية تطبيقية لأنه يعتمد على تطبيق بعض التصورات اللسانية، ومحاكاة بعض الأنشطة اللغوية، وتشغيلها ضمن برامج حاسوبية.
9. إن البرامج الحاسوبية التي حاولت صياغتها تعمل على تكوين الجمل العربية بطريقة آلية، فهي تنطلق مما زود به برنامج برولوغ من وقائع وقواعد تتعلق بالمطابقة بين وحدات الجمل العربية، فيقوم البرنامج بتكوين جمل صحيحة

التركيب سليمة المعنى بعد أن يطلب منه ذلك، وإن بدا لي أن بعض تلك الجمل ليس مقبولا عند المتكلمين العرب، من أجل ذلك آمل في بحوث مستقبلية تطوير هذه البرامج بإضافة بعض القيود التي تتعلق بالجوانب الثقافية والاعتقادية.

10. إن دراسة ظاهرة المطابقة بين الفعل وفاعله في الجملة الفعلية، وبين المبتدأ والفعل الواقع خبرا له في الجملة الاسمية الكبرى وتشغيلها حاسوبيا، تؤكد نظرية النحاة التي تتضمن عدم تماثل هاتين البنيتين من حيث الشكل، على عكس ما يعتقده بعض النحاة القدماء وبعض اللسانيين المحدثين من أن الفاعل يمكن تقديمه عن فعله من غير تغيير في بنية الجملة.

وفي الأخير، أرجو أن أكون قد وفقت إلى دراسة هذا الموضوع، وإلى تناول أبرز مكوناته، وآمل أن أثريه في بحوث مستقبلية من خلال تدعيم جانبه التطبيقي الخاص بالبرمجة الحاسوبية حتى أحقق الغرض الأسمى من هذه الدراسة: وهو المساهمة –بجهدي المتواضع وببضاعتي المزجاة– في إقامة بحث يتناول قضية معاصرة باللغة العربية، وتمكين هذه اللغة من أن تكون حاضرة جنبا إلى جنب مع سائر اللغات الحية المعاصرة من خلال مختلف التطبيقات الحاسوبية التي تنجز بها.

والله من وراء القصد، وهو يهدي السبيل، وصلى الله على سيدنا محمد وآله.

مسرد المصطلحات

لإمصطإح بلإلإة الإلإليلإمة	لالمصطلاح باللغة لافرنسيلا	لامصطلاح باللغة العربيلا
Automation	Automatisation	أتمتة
Choice	Choix	لالاتيار
Strategic	Stratégique	استرا تيجي
Using	Utilisation	استعمال
Finite-state automaton	Automate à état fini	لاوا ليةلالالالات لامتلاهية
Programming	Programmation	لالامجة
Softaware	Logiciel	برمجية
Program	Programme	برنامج
Rhetorical	Rhétorique	خطابي (بلاغي)
Rhetoric structuration	Structuration rhétorique	البنمنة الخطابية
Surface Realization	Réalisation de surface	تحقة السطح
Natural Language Analysis	Analyse Automatique de Textes	تحليل الآ لي للنصوص
Pragmatic	Pragmatique	لالأولآم
Compilation	Compilation	لآلآلالآ
Synthesis	Synthèse	ترلآمب
Syntactic	Syntaxique	تركيبي
Symbolization	Symbolisation	ترمملآ
Serial	Sériel	تسلسلي (لآموذج)
Implementation	Implémentation	تشغيل لآي
Encoding	Encodage	تشفير
Interface	Interface	لآصافح
Planning	Planification	تصميم
Application	Application	تطبيق
Simulation	Simulation	تقييس (محاكاة)
Tactical	Tactique	تكتمكي
Adaptation	Adaptation	تكييف

Stratificationnel	Stratificationnelle	لالآمـلآيلاً أو لالآمـلا(لالالالآ)
Natural Language Generation	Génération Automatique de Textes	الآتوليد الآ ليـ للنصوص
Theme	Thème	لآي (موضولاً)
Sentence	Phrase	جملة
Software	Software	جهاز برمجي
Hardware	Hardware	لآهلآلاً مالآمـ
Computation	Computation	لآولآبة
Algorithm	Algorithme	خوارزم
Algorithmic	Algorithmique	لآوارزميلاً
Artificiel Intelligence	Intelligence Artificielle	لالالالا لالاصلآذلآلآي
Rheme	Rhème	ريم
Discriminant network	Réseau discriminant	لالالالة مميزة
Flexion	Flexion	لالافة
Morphological	Morphologique	صرفي
Formalism	Formalisme	صورانمة (ألالالا)
Formalization	Formalisation	لاونة
Paraphrases	Paraphrases	لايـ لالا لملاعددة
Cognitive sciences	Sciences cognitives	العلوم الإدراكية أو المللإلإيلإ
Neuroscience	Neurosciences	العلوم العصبية
Communicational intent	Intention communicationnelle	غ لإلإ تواصلي
Ambiguity	Ambiguïté	غموض
Agent	Agent	فاعلإ
Clause	Proposition	قضية
Database	Base de données	قاعدة بيانات (ملإلإيات)
Knowledge base	Base de connaissances	لإلإلإلإلإ الإلإ
Grammaticality	Grammaticalité	قولإلإيـة
Psycholinguistics	Psycholinguistique	لسانيات نفسية
Language	Language, langue	لغة
Lemme	Lemma	لمة
Lexeme	Lexème	لإوسيم
Synthesizer	Synthétiseur	لإلإ لإف
Semantic content	Contenu sémantique	محتوىلالالي
Predicate	Prédicat	محمولا

Discourse specify	Specificateur du discours	لالا لاص لاخطلالا
Stereotyped diagram	Schémas stéréotypés	لامخططات المقولاة
Integrated (approach)	Intégrée (approche)	ةجمدم(لالالالة)
Phrase	Syntagme	لاللاب
Discourse planning	Planificateur du discours	لا لالة لالالاطاب
Microprocessor	Microprocesseur	لاالج صغ لام
Natural Language Processing	Traitement Automatique des Langues Naturelles	المعالجة الآلية للغات الطبيعية
Lexicon	Lexique	ملالام
Lexicalization	Lexicalisation	معجمة
Humain cognition	Cognition humaine	المعرفة الإلالانملا
Architecture	Architecture	معمارمة
Meaning	Sens	مـلإنم
Conceptualization	Conceptualisation	مفهمة
Constituent	Constituant	لإكلإن
First-order logic	Logique de 1er ordre	لألألأ لألألألألإلأ لألألإ
Module, modular	Module, modulaire	لأنظومة، لأنظ ومم
Pattern, model	Modèle	منوال أو نموذلأ
Argument	Argument	موضوع
Object	Objet	لأولأوع
Discourse Grammar	Grammaire de Discours	نحو الخلألألأ
Text	Texte	لألأ
Modality	Modalité	نلأط
Expert system	Système expert	نظام خبير
Modelization	Modélisation	نمذجلأ
Data structre	Strucutre de données	هيكل بيانات

قائمة المصادر والمراجع

- **باللغة العربية**

- القرآن الكريم.

1. إدريس (سهيل)، المنهل، دار الآداب، بيروت، ط.11، 2003.
2. أدهم (سامي)، إبستمولوجيا المعنى والوجود ونقد التطورية، مركز الإنماء القومي.
3. استيتية (سمير شريف)، اللسانيات: المجال والوظيفة والمنهج، عالم الكتب الحديث، الأردن، ط. 1، 2005.
4. الأسدي (حسن عبد الغني جواد)، مفهوم الجملة عند سيبويه، دار الكتب العلمية، بيروت، ط.1، 2007.
5. أنسو (فرانسوا)، « من **فونيومن Von Neumann** إلى المعالجات الصغرية الفائقة »، ما التكنولوجيا ؟ تر. محمد نايت الحاج وعبد الهادي إدريسي، المجلس الأعلى للثقافة، القاهرة، 2005.
6. الأنصاري (ابن هشام)، شرح قطر الندى وبل الصدى، تحقيق محمد محيي الدين عبد الحميد، مطبعة السعادة، مصر، ط.11، 3963.
7. الأنصاري (ابن هشام)، مغني اللبيب عن كتب الأعاريب، ت. محمد محيي الدين عبد الحميد، المكتبة العصرية، بيروت، 1995.
8. أنيس (إبراهيم)، من أسرار اللغة، مكتبة الأنجلو المصرية، القاهرة، ط.3، 1966.

9. ابن جني (أبو الفتح)، الخصائص، ت. محمد علي النجار، المكتبة العلمية، د.ط، د.ت.

10. ابن عقيل، شرح ابن عقيل، ت. حنا الفاخوري، دار الجيل، بيروت، ط.5، 7997.

11. ابن فارس (أبو الحسين)، الصاحبي في فقه اللغة العربية ومسائلها وسنن العرب في كلامها، ت. عمر فاروق الطباع، مكتبة المعارف، بيروت، ط.1، 1993.

12. ابن يعيش، شرح المفصل، تحقيق: أحمد السيد أحمد، المكتبة التوفيقية، القاهرة، د.ط، د.ت.

33. بوول (ج.)، النظرية النحوية، تر. مرتضى جواد باقر، المنظمة العربية للترجمة، بيروت، ط. 1، 2009.

14. الحاج صالح (عبد الرحمن)، بحوث ودراسات في اللسانيات العربية، موفم للنشر، الجزائر، د.ط، 2007.

15. الدجني (فتحي عبد الفتاح)، الجملة النحوية نشأة وتطورا وإعرابا، مكتبة الفلاح، الكويت، ط. 2، 7987.

16. ديتربونتنج (كارل)، المدخل إلى علم اللغة، تر. حسن بحيرى، مؤسسة المختار، القاهرة، ط.1، 2003.

77. زكريا (ميشال)، الألسنية التوليدية والتحويلية وقواعد اللغة العربية، المؤسسة الجامعية، بيروت، ط. 2، 1986.

18. الزناد (الأزهر)، نظريات لسانية عرفنية، منشورات الاختلاف، الجزائر، ط.1، 2010.

19. عبد الرحمن (طه)، اللسان والميزان، المركز الثقافي العربي، الدار البيضاء، ط.1، 1995.

20. عبد اللطيف (محمد حماسة)، العلامة الإعرابية في الجملة بين القديم والحديث، دار غريب، القاهرة، د.ط، 2001.

21. علوي (حافظ إسماعيلي) والملاخ (امحمد)، قضايا إبستمولوجية في اللسانيات، منشورات الاختلاف، الجزائر، ط. 1، 2009.

22. العناتي (وليد) والجبر (خالد)، دليل الباحث إلى اللسانيات الحاسوبية العربية، دار جرير، عمان، ط. 1، 2007.

23. الفاسي الفهري (عبد القادر)، اللسانيات واللغة العربية، دار توبقال، المغرب، ط.1، 1985.

24. ماكارثي (ميشيل)، قضايا في علم اللغة التطبيقي، تر. عبد الجواد توفيق محمود، المجلس الأعلى للثقافة، القاهرة، ط.1، 2005.

25. المبرد، المقتضب، تحقيق محمد عبد الخالق، عالم الكتب، بيروت، د، ط.

26. المتوكل (أحمد)، الوظائف التداولية للغة العربية، دار الثقافة، الدار البيضاء، ط.1، 1985.

27. المخزومي (مهدي)، مدرسة الكوفة ومنهجها في دراسة اللغة والنحو، مطبعة مصطفى الحلبي، مصر، ط.2، 1958.

28. المخزومي (مهدي)، في النحو العربي، نقد و توجيه، دار الرائد العربي، بيروت،ط. 2، 1986.

29. معجم الحاسبات، منشورات مجمع اللغة العربية، القاهرة، ط. 2، 1995.

30. منصوري (ميلود)، دلالات التراكيب في نحو الجملة، دار أم الكتاب، مستغانم، ط.1، 2900.

11. الموسى (نهاد)، «حصاد القرن في اللسانيات» ضمن: حصاد القرن: المنجزات العلمية والإنسانية في القرن العشرين، مؤسسة عبد الحميد شومان، الأردن، د.ت.

32. الموسى (نهاد)، العربية: نحو توصيف جديد في ضوء اللسانيات الحاسوبية، المؤسسة العربية للدراسات والنشر، بيروت، ط.1، 2000.

33. نحلة (محمود أحمد)، مدخل إلى دراسة الجملة العربية، دار النهضة العربية، بيروت، 1988.

34. الواسطي (سلمان داود)، "التفاعل بين الإنسان والآلة في الترجمة الحاسوبية"، مجلة التعريب، المركز العربي للتعريب والترجمة، دمشق، ديسمبر 2000، العدد 20.

35. الوعر (مازن)، قضايا أساسية في علم اللسانيات الحديثة، دار طلاس، دمشق، ط.1، 1988.

- **باللغة الأجنبية**

36. Abeillé (Anne), Les grmmmii sss’’’unifia æaa a aaaaaaaa
37. Abeillé (Anne), Les nouvelles syntaxes : gmmmii d’uniiittt ion et analyse du Français, Armand Colin, Paris, 1993.
38. Alario (François-Xavier) et *al*., Psycholinguistique cognitive, De Boeck, Bruxelles, 1re éd., 2004.
39. Al Daimi (J. Khalid) and Abdel-Amir (Maha A.), « The syntactic analysis of arabic by machine », Computers and Humanities, vol. 28, n° 1, Springer, Netherlands, 1994.
40. Bach (Emmon), Introduction aux grammaires transformationnelles, trad. Robert Sctrick, Armand Colin, Paris, 1973.
41. Balicco (Laurence), *Génération de répliques en Français dans une interface Homme-Machine en langue naturelle*, thèse de doctorat, Université de Grenoble, 1993.
42. Bolshakov (I.) & Gelbukh (A.), Computational linguistics: models, resources, applications, Instituto Politecnico Nacional, Mexico, 2004.
43. Bouillon (Pierrette) & *al*., Traitement automatique des langues naturelles, Aupelf-Uref – Ed. Duculot, Belgique, 1998.
44. Covington (A. Michael), Natural language processing for Prolog programmers, Prentice-Hall, New Jersey, 2013.
45. Cuvillier (Armand), Nouveau vocabulaire philosophique, Armand Colin, Paris, 1956.
46. Danlos (Laurence), Génération automatique de textes en langues naturelles, Masson, Paris, 1985.

47. Delbecque (Nicole), Linguistique cognitive, De Boeck, Bruxelles, 2e éd. 2006.
48. Ditters (Everhard), ‘TTh Desiii ption o oo drrn Standadd Abbi yyntax in Terms of Functions and Traits’’ *in* Langues et Littératures du Monde Arabe, 2 (2001), pp. 115-151.
49. Dubois (Jean) et *al*., Dictionnaire de linguistique et des sciences du langage, Larousse, Paris, 1999.
50. Dufaye (Lionel), Théorie des opérations énonciatives et modélisation, éd. Ophrys, 2009.
51. El Kassas (Dina), *Une étude contrastive de l’arabe et du français dans une perspective de génération multilingue*, thèse de Doctorat, Université Paris7, 2005.
52. Fayol (Michel), Production du langage, Lavoisier, 2002.
53. Ferrand (Ludovic), « La production du langage : un vu d’nnsemble », *in* Psychologie française, mars 2001, tome 46 – n°1, PUG.
54. Fromkin (V.) & *al*., An introduction to language, Wadsworth, USA, 7th ed.
55. Fuchs (C.) & *al*., Linguistique et traitement automatique des langues, Hachette, Paris, 1993.
56. Fuchs (C.) & Le Goffic (P.), Les linguistiques contemporaines: repères théoriques, Hachette, Paris, 1992.
57. Gardent (Claire) & sss chung KKii n)) Thhhniques d’nnalys tt de génération pour la langue naturelle, Ed. Adosa, 1994.
58. Gineste (M-D) & Le Ny (J-F), Psychologie cognitive du langage, Dunod, Paris, 2002.
59. Lazar (Judith), La science de la communication, Editions Dahlab, Alger, 2e éd., 1993.
60. Lemaire (Patrick), Psychologie cognitive, De Boeck, Paris, 1999.

61. Léon (Jacqueline), « De la traduction automatique à la linguistique computationnelle. Contribution à une chronologie des années 1959-1965 », Traitement Automatique des Langues N° spécial trentenaire, 1992, vol.33, n° 1-2 :25-44.
62. Le Robert : Dictionnaire Français-Anglais, Anglais-Français, Paris, 6e éd., 2002.
63. Ligozat (Gérard), Représentation des connaissances et linguistique, Armand Colin, Paris, 1994.
64. Malmkjær (K.), The linguistics encyclopedia, Routledge, London, 1st ed., 1991.
65. lll ’u u IIgor A.) & *al*., Introduction à la lexicologie explicative et combinatoire, Duculot, Belgique, 1995.
66. Mitkov (R.), The Oxford Handbook of computational linguistics, Oxford University Press, New York, 1st published, 2003.
67. Morvan (P.), Ditt ionnii ddd’’’infom tt tt t ooœoo rrr rrr
68. Mounin (George), Dictionnaire de linguistique, PUF, Paris, 1974.
69. Mounin (George), « Le modèle en linguistique », tome XV, Encyclopædia Universalis, Paris, 2002.
70. Neveu (Franck), Lexique des notions linguistiques, Nathan, Paris, 2000.
71. Pierrel (J.M.), Ingénierie de la langue, Hermès, Paris, 2000.
72. Polguère (A.), *Structuration et mise en jeu procédurale d'un modèle linguistique déclaratif dans un cadre de génération de texte*, thèse de doctorat, Université de Montréal, 1990.
73. Ponton (Claude), *Génération automatique de textes en langue naturelle : essai de définition d'un système noyau*, thèse de doctorat, Grenoble3, 1996.

74. Pylyshyn (Zenon W.), « Computation and cognition: issues in the foundations of cognitive science », The Behavioral and Brain Sciences, USA, 1980, N°3, 111-169.
75. Rastier (François) & *al*,, mmmn mmn n nmmn : de la linguistique à mmmnnnnmmn rrr rr, 1994.
76. Ratté (Sylvie), *Interprétations des structures syntaxiques : une analyse computationnelle de la structure des évènements*, thèse de doctorat, Université de Montréal, 1995.
77. Rey-Debove (Josette) & Rey (Alain), Le nouveau Petit Robert, Dictionnaires Le Robert, 1995.
78. Ruwet (Nicolas), Introduction à la grammaire générative, Plon, Paris, 1re éd., 1968.
79. bbbah GGaaaad ↄ : 2. Processus de compréhension, Hermes, Paris, 1989.
80. Stehlé (Jean-Luc) et Hochard (Pierre), Ordinateurs et langages, Ellipses, Paris, 1989.
81. Tssnieees LLuii en, , , ,, , ,,, Paris, 1953.
82. Théodule (M.-L.) & Gruszow (S.), « Quelles langues vont disparaître ? », La Recherche, n° 429, Avril 2009.
83. Trask (R.L.), A dictionary of grammatical terms in linguistics, Graham McFee, 1996, « Peters & Ritchie results ».
84. Vander Linden (K.), Speech and language processing: An introduction to speech recognition, computational linguistics and Natural language processing, Prentice Hall, New Jersey, 1998.
85. Yousfi (Abdellah), Traitement automatique de la langue (texte et parole), Bouregreg, Rabat, 1re éd., 2006.

- **المراجع الإلكترونية**

86. غازي (عز الدين)، "اللسانيات الحاسوبية واللغة العربية"، الحوار المتمدن، مجلة إلكترونية، العدد 3639، 2006، http//www.alhewar.org

87. "اللغة العربية": اللغة_العربية/http//ar.wikipedia.org/wiki

88. Abeillé (Anne), « Traitement automatique des langues » *in* Encyclopædia Universalis, Paris, 2002. CD-ROM.

89. Bies (Ann) and Maamouri (Mohamed), PENN ARABIC TREEBANK GUIDELINES, 2003, http//www.ircs.upenn.edu/arabic/Jan03release/guidelines-TB-1-28-03.pdf

90. Glossaire, Service ATO (UQAM) et EBSI (Université de Montréal). http://www.ling.uqam.ca.

91. Hajic (Jan) et *al*., Prague Arabic Dependency Treebank: Development in Data and Tools, http://ufal.mff.cuni.cz/padt/PADT_1.0/docs/index.html.

92. Monin (Jean-François), « Programmation » *in* Encyclopædia Universalis, Paris, 2002. CD-ROM.

93. Pastre (Dominique), Cours de Prolog, 1990/2000, http://www.math-info.univ-paris5.fr/~pastre/prolog/cours.pdf.

94. Tll li (,,,, Introdutt ion au TALN tt l’ingnnirri linguistique, Université Lille3, http:// www.grappa.univ-lille3.fr/polys/info_ling/index.html.

95. Vuillaume (Marcel), Variation sur la référence verbale, google book.

فهرس الأشكال والجداول

الأشكال	الصفحة	الجداول	الصفحة
الشكل1: الكلام والجملة عند ابن جني	10	الجدول1: الكلام المستغني عند سيبويه	4
الشكل2: نموذج من مشجر تعلقي	36	الجدول2: تصور الكلام عند ابن فارس	7
الشكل3: بنية المركب الاسمي	39	الجدول3: تكافؤ البنى التركيبية العربية	18
الشكل4: بنية المركب الفعلي	40	الجدول4: طرق الارتباط بين علوم اللغة وعلوم الحاسوب	49
الشكل5: التقابل بين اللسانيات المعلومية والمعلوميات اللسانية	54	الجدول5: مستويات التجريد في الدراسات الحاسوبية	64
الشكل6: نموذج الحالات المتناهية	58	الجدول6: توليد/فهم في مقابل تركيب/تحليل	129
الشكل7: خطوات المعالجة وأدواتها	33	الجدول7: نماذج لأنظمة التوليد	152-353
الشكل8: محاور البحث في المعالجة الآلية للغة	94		
الشكل9: نموذج إنتاج الكلام المنطوق حسب لفيلت Levelt	111		
الشكل10: نموذج بينة شبكة عصبونات	124		
الشكل11: معمارية نظام التوليد الآلي للنصوص	334		
الشكل12: الحوار بين الإنسان والآلة	777		
الشكل13: المولد الآلي من وجهة تداولية	338		
الشكل14: مكونات المولد ووظائفها	142		
الشكل15: ملف كتابة البرنامج ذو الامتداد *pl*	778		
الشكل16: البرنامج المنفذ للبرولوغ	779		

الفهرس العام للموضوعات

شكر وتقدير

إهداء

مـــقدمـــة..... أ–ط

مفـــــاهيم أوليـــــة..... 14

الفصل الأول: مفاهيم أساسية 15

1. الجملة عند النحاة..... 15

1.1. الجملة عند سيبويه..... 15

2.1. أولية المصطلح..... 18

3.1. الجملة عند ابن فارس وابن جني 19

2. الجملة في الدراسات الحديثة..... 24

1.2. الجملة عند الدارسين العرب..... 24

2.2. الجملة عند الدارسين الغربيين 33

3. المكونات الخطية للجملة..... 37

1.3. الكلم المولدة في الجملة الفعلية..... 37

2.3. الكلم المولدة في الجملة الاسمية 40

4. اللسان العربي 41

1.4. الخريطة الجغرافية 41

2.4. النحو الصوري للغة العربية 43

الدراسة اللسانية الحاسوبية..... 51

الفصل الثاني: الدراسة اللسانية الحاسوبية..... 52

1. اللسانيات الحاسوبية: المفهوم والمنهج..... 52

1.1. مشكل المصطلح والترجمة..... 53

2.1. اتصال اللسانيات بالنظريات الحاسوبية..... 64

3.1. منهج اللسانيات الحاسوبية 68

2. المعالجة الآلية للغات الطبيعية..... 73

1.2. المفاهيم الأساسية 76

2.2. صعوبات المعالجة 79

3.2. مقاربات المعالجة... 82

4.2. موارد المعالجة..... 90

5.2. أدوات المعالجة..... 93

3. محاور البحث في "المعالجة الآلية للغة" 94

الفصل الثالث: نماذج التوليد اللسانية واللسانية النفسية 100
1. النماذج اللسانية 101
1.1. الأنحاء التوليدية 104
2.1. الأنحاء المعرفية والوظيفية المعرفية 107
2. النماذج اللسانية النفسية 111
1.2. معمارية إنتاج الكلام 112
2.2. مراحل إنتاج الكلام 116
3.2. نمذجة إنتاج الكلام 123
التوليد الآلي للجمل 127
الفصل الرابع: التوليد الآلي للجمل 128
1. التوصيف 128
2. الأبحاث المنجزة 136
3. مكونات المولد ووظائفها 140
1.3. مكونات المولد 140
2.3. وظائف مكونات المولد 142
4. نماذج من أنظمة التوليد المنجزة 149
1.4. نموذج من مولدات اللغة العربية 154
2.4. ملاحظات حول أنظمة التوليد 156
5. تقنيات التوليد 157
6. خوارزمات التوليد 163
التوليد الآلي للجمل العربية 165
الفصل الخامس: التوليد الآلي للجمل العربية 166
1. النموذج المركبي للجملة العربية 169
2. لغة برولوغ 175
3. نحو العبارات المحددة للجملة العربية 181
4. توليد الجمل 183
1.4. توليد الجمل الاسمية 184
2.4. توليد الجمل الفعلية 203
خـــــاتمـــة 210
خـــاتـــــمـــــة 211
مسرد المصطلحات 215
قائمة المصادر والمراجع 218
فهرس الأشكال والجداول 227
الفهرس العام للموضوعات 228

الموضوع: *توليد الجمل في اللسان العربي – دراسة لسانية حاسوبية –*

الملخص:

في هذه الرسالة، قمت بدراسة عملية توليد الجمل العربية دراسة لسانية حاسوبية. من أجل ذلك، اعتمدت على منهجية تتضمن تعريفا ببعض المفاهيم الأولية كالجملة واللسان العربي، لأتناول بعدها أسس الدراسة الحاسوبية وموضوعها ومنهجها. وبما أن عملية التوليد هي عملية لسانية ولسانية نفسية كذلك، خصصت جانبا من هذه الرسالة لعرض نماذج التوليد من المنظورين اللساني واللساني النفسي. والملاحظ أن هذه الجوانب: اللسانية الحاسوبية واللسانية النفسية تتقاطع فيما يسمى بالتوليد الآلي للنصوص حيث اقتضى الأمر أن أتطرق إلى عناصره. وقد ختمت هذه الرسالة بتطبيق يتضمن إنشاء برامج حاسوبية لتوليد الجمل العربية باستخدام لغة البرولوغ.

الكلمات المفتاحية: توليد الجمل، اللسان العربي، اللسانيات الحاسوبية، لغة برولوغ.

Thème : *La génération de phrases en langue arabe -étude linguistique informatique-*

Résumé :

Dans cette thèse, j'ai essayé d'étudier la génération des phrases arabes dans le domaine de la linguistique informatique. Pour cela, j'ai procédé par la méthodologie suivante : primo, la définition de quelques concepts préliminaires (phrase, langue arabe), secundo, la détermination du domaine de la linguistique informatique. Et étant donné que la génération de phrases est un processus psycholinguistique, j'ai exposé les modèles de cette dernière à côté de celles de la linguistique. Il est aussi clair que la génération automatique des phrases est un domaine mixte de la linguistique informatique et de la psycholinguistique pour laquelle j'ai consacrée tout un chapitre. Finalement, l'application de cette étude consiste à construire des programmes informatiques pour la génération des phrases arabes en utilisant le langage Prolog.

Mots-clés : génération de phrase, langue arabe, linguistique informatique, langage Prolog.

Theme : *Generating sentences in Arabic language - computational linguistics study-*

Abstract :

The nature of this work requires a set of questions that are raised among a general problematic whose answer comes from components of this thesis. This research encompasses an introduction and five chapters. The first chapter deals with some concepts such as, sentence and language where different definitions are presented. The second chapter spells out basics of computational and linguistic study .The third chapter revolves around a discussion of generating sentences with regard to different linguistic and psycholinguistic theories. The fourth chapter talks about Automatic generation. The last chapter entails a practical survey hinging upon a program written by PROLOG in order to form Arabic sentences automatically.

Keywords: generating sentences, Arabic language, computational linguistics, Prolog language.

Algerian Democratic Popular Republic
Minister of Higher Education and scientific Research
Abou Bekr Belkaid University – Tlemcen
Literatures and Languages School Department of Arabic language and literature

Abstract of PhD Outline

Tittle:

Generating Sentences in Arabic Language

"A Computational Linguistics Study"

Preparing by:

- BABA AHMED REDA

Director of Research:

- RITRI SIDI MOHAMMED

University Year 2013/2014

Introduction

Modern Researches in computational fields have achieved myriad goals in terms of information processing in general and linguistics in particular. Programmes and systems used by computer users have spread out tremendously providing many services such as text treatment, correcting grammatical and orthographic mistakes, translation, communication throughout internet, and other programmes related to language processing.

The fore mentioned programmes are supplied with systems to produce and analyze language i.e., to generate different linguistics data such as texts, paragraphs, sentences, and linguistic units relying on rules of data and representations of each datum. Furthermore they analyze different linguistic data and recognize coded data. Many languages like English and French were interested in deep study to design sophisticated computational programmes to ease the task for users in terms of linguistic activities and interaction.

Arabic language as many other languages, represents significant functions and uses through its different varieties. Not only in cultural, religious, and literary side, but also in scientific, economic and social sides in the realm of technology. For this reason Arabic language processing is a challenging task for researchers. As mentioned before, programmes of natural language programming hinges on two psycholinguistic factors; generation and aa a sss iii i iiii i i ggggggggggggg ***Generating Sentences in Arabic Language" (A Computational and Linguistic Study)*** was chosen to deal with the generation of Arabic sentences from the computational and linguistic sides.

The nature of this work requires a set of questions that are raised among a general problematic whose answer comes from components of this thesis. The questions are: How can the Arab speaker produce and generate sentences? Is it possible to do this automatically? The answer to these questions requires to scaffold on a set specialties such as Arabic grammar, Linguistics, Psycholinguistics, and Computational Linguistics.

It is undoubtedly that there is a set of things which spurred me to tackle this work. First the competence of the speaker who generates meaningful sentences built together appropriately to form texts and speech, second I was fascinated by the way researchers describe this operation and their different views in interpreting and displaying its components, furthermore computational study for generating sentences is necessary to design programs such as automatic translation which necessitates a deep description of Arabic texts and sentences, in addition to theoretical systems in order to formulate ways of generating sentences, because computational study of generating sentences requires a set of theories such

as linguistics and psycholinguistics. For this reason different hypotheses were suggested to interact with different communicative and computational theories.

In a broad of view this work aims to use a set of computational linguistic basics to design a computational program able to generate sentences. To this effect a practical survey is made. It entails building Arabic sentences via PROLOG (one of the language of programmation), so this study is very limited in the sense that it cannot present a complete conception of a very developed program which can generate all types of Arabic sentences.

Methodology

To design the survey of this research, this part outlines the items intended to answer the questions posed in the introduction. It cast light on fundamental concepts (concept of sentence and generation) used in the research ,then implements them in the study, then describes Arabic language and finally formalizes it so as to function automatically.

This research encompasses an introduction and five chapters. The first chapter deals with some concepts such as, sentence and language where different definitions are presented. The second chapter spells out basics of computational and linguistic study .The third chapter revolves around a discussion of generating sentences with regard to different linguistic and psycholinguistic theories. The fourth chapter talks about Automatic generation. The last chapter entails a practical survey hinging upon a program written by PROLOG in order to form Arabic sentences automatically. As for the conclusion, it encompasses most significant results of the study, a list of references, a glossary at the end of the research, and an annex of computational programs and indexes devoted to graphs, charts and titles.

Hardships

The hardships faced in this research can be summarized in the multiple theoretical frameworks which dealt with the sentence and sometimes are opposed as do formalistic, functional, and cognitive linguistics in addition to different specialties which were used as a scaffolding in this research mainly linguistics and computer science which was very difficult to join them.

Conclusion

This work tries to cast light upon different concepts related to the topic, for instance, sentence, language and generation. It also displays elements of computational and linguistic study, its notion, methodology, and its purpose with regard to the study of language. It also aims to explain the process of generating sentences which embraces different theories taken from formalism, functionalism, and cognitivism. It spells out basics of automatic generation

of sentences and ends with a practical survey containing designing process of computational programs of generating sentences of Arabic language using PROLOG after displaying a set of automatic generators considered as a hinge to this operation.

The results of this work can be summarized as follows:

1. The concepts of sentence and speech are very difficult to determine ,however many grammarians and modern researchers have put borders between these concepts such as relating the notion of sentence to speech and its skeleton, and the notion of speech to the communication process.
2. Samples that explain the process of generating sentences in psycholinguistic research are sequential or parallel, where the first one has developed significantly compared to the parallel one which still needs some research. The sequential one relies on hypotheses of generative grammar associated new conceptual cognition
3. Arabic literature which dealt with computational linguistics has associated linguistics to computer sciences though there are very delicate areas and is difficult to separate them in computational linguistics ,in addition to the problem of terminology still remains unsolved i.e., many terms for one concept for instance الهندسة اللسانية (linguistic engineering) تكنولوجيا اللسان ، (technology of the language) in this respect researcher should endeavour to unravel the underbelly of this terms.
4. I suggested a new term translated from the French term (linguistique informatique) so as to avoid ambiguity therefore to determine a particular field of study
5. A lot of imperfections have been noticed by researchers regarding the current generating systems. Perhaps they are related to a specific application and are certainly used with special artificial languages that cannot be generalized on other languages, more simply put their texts have special forms, so the linguistic sample is imperfect.
6. The computational programs that I have tried to design, intends to form Arabic sentences automatically. It uses PROLOG as a support because it works with rules of similarities between units of Arabic sentences in the sense that it forms sentences that are correct in meaning when asked. My future perspective is to develop this program and enable cultural and religious sides. At the end despite the great effort devoted to this work in dealing with most of its aspects, much work still needs to be done to enrich it mainly in the practical survey which concerns computational programmation, so as to contribute to enable Arabic to get a share among languages of technology.

..جمهورية الجزائرية الي.ق .ي. ـية
وزارة الـ..مـ. ... .مـ والبـ .ث ا لعلمي
جامعة .ي .كر بلقام. - تلمسان
لآملآ لآلآداب وا للغات قسلآلألأ ولآلآ لآلآربي

ملخص رسالة الدكتوراه

في اللسانيات التطبيقية بعنوان:

توليد الجمل في اللسان العربي

(دراسة لسانية حاسوبية)

إعداد الطالب:
- بابا أحمد رضا

إشراف الأستاذ الدكتور:
- غيثري سيدي محمد

السنة الجامعية 2014/2013

إن التفكير في الجملة وما تحتوي عليه من عناصر وما تقابلها من دلالات تركيبية من أول ما شد انتباه الدراسين العرب؛ فالمتتبع لما ورد في كتب النحو العربي يجد اختلافا بين النحاة في تحديد الجملة، بل يجد تضاربا حول اتخاذها مصطلحا للدلالة على مفهوم علمي معين. في أقدم وأكمل كتاب نحوي وصل إلينا، نعني كتاب سيبويه، نلفيه يستخدم في كثير من المواضع مصطلحا قريبا من مصطلح الجملة وهو الكلام المستغني.

أما عن مصطلح الجملة فيرجح عبد الرحمن الحاج صالح بأن الأخفش تلميذ سيبويه وراوية كتابه وأستاذ المازني قد يكون أول من ابتدع مصطلح الجملة مقابلا للكلام المستغني لأنه أشار في كتابه "القوافي" إلى أبرز مكون لمفهوم الجملة ألا وهو الفائدة، ويظن أنه انتقل عبر المازني إلى المبرد.

يبدو التفريق بين الكلام والجملة أكثر وضوحا عند ابن هشام الأنصاري حيث يعرف الكلام بأنه القول المقصود المفيد الذي يحسن السكوت عليه، أما الجملة فهي عبارة عن الفعل وفاعله، والمبتدأ والخبر، وما كان بمنزلة أحدهما، فلا يشترط في الجملة أن تكون مفيدة، والدليل على ذلك قول النحاة: جملة الشرط، جواب الشرط، جملة الصلة، وهذه كلها ليست كلاما لأنها ليست مفيدة.

من جهة أخرى، يتعدد مفهوم الجملة عند علماء اللغة العربية المحدثين بسبب انتمائهم إلى مدارس ومذاهب لغوية شتى، بعضها متأثر بالدراسات العربية في أطوارها المتتالية، وبعضها متأثر بنظريات لغوية غربية، وبعضها مازج بين القبيلين، وقد يترتب عن ذلك اختلاف القواعد والأحكام اللغوية الخاصة بالجملة، فهناك من اللغويين العرب من يرى أن الجملة قول مركب مفيد دال على معنى يحسن السكوت عليه، فلا يخرجون بذلك عما درج عليه النحاة عامة والمتأخرون على وجه الخصوص.

وعلى العموم، اختلف علماء اللغة في مسائل شتى حول الجملة النحوية وتحديد مفهومها وتسميتها وأقسامها وأنواعها، إلا أنهم اتفقوا على أن الجملة لا تتألف من جزء واحد، إذ لا بد أن

يشتملا على كلمتين على الأقل، وكل منهما يتوفر على مسند ومسند إليه، وبغيرهما لا يستقيم كلام ولا تتحقق فائدة.

لقد ركز النحو التقليدي –الفرنسي خاصة والأوروبي عموما– على اعتبار الجملة تجمُّعا للكلمات من أجل تشكيل معنى تام، وهي تتميز عن القضية بحيث إن الجملة قد تتضمن عدة قضايا (في الجملة المركبة مثلا). ولكن هذا التعريف في نظر بعض الدارسين غير عملي؛ إذ يمكن أن يضمَّن المعنى في جملتين أو أكثر، كما أن المعنى قد يكون تاما دون أن يكون هناك بناء سليم للجملة.

ومن جهة نظر بنوية، تعتبر الجملة قطعة من السلسلة الكلامية مستقلة من الناحية التركيبية؛ أي هي شكل لساني مستقل غير متضمَّن في شكل لساني أكبر بموجب بنية نحوية، بعبارة أخرى: هي مكوَّن لا ينضوي في مكوَّن أكثر امتدادا، وهي وحدة كبرى للوصف النحوي، لا تنتمي إلى صنف توزيعي، وبالتالي لا يمكن أن يتم الاستبدال بينها وبين مكونات أخرى.

أما من وجهة نظر التوليديين، فإن الجملة هي أولية قاعدية (أكسيوم) تتفرع عنها سلسلة من الرموز بواسطة قواعد تركيبية. الجملة P هي التي تتفرع من القاعدة الأولى: $P \rightarrow SV + SN$ حيث إن P هو الرمز الأولي، وSV وSN هما على التوالي: المركبان الفعلي والاسمي.

ترتبط الجملة في بنائها بفكرة العامل ارتباطا وثيقا؛ فهو الذي يقوم بوصل مكوناتها ويجعلها نسيجا متلاحما، ويشكل معها وحدة معنوية وبناء متكاملا. وقد ركز النحاة كثيرا على فكرة العامل، وبينوا أهميتها في النحو العربي، ويذكر **الأسدي** بأن سيبويه يتناولها باستمرار، وهو يقصد بها إمكانية توليد المجالات أو إفراغها أو إنشائها لتشغلها المقولات الاسمية وترتبط بالعامل، فتدل على الوظائف النحوية. أما الجملة عنده فهي الهيأة الحاصلة من العمل، وهو الهيكل الذي يشمل مختلف الوظائف النحوية.

يتكلم بالعربية اليوم أكثر من 422 مليون شخص في 22 دولة على الأقل:

- الجزيرة العربية: السعودية، البحرين، الإمارات المتحدة، عمان، قطر، اليمن؛

- الشرق الأوسط: العراق، الأردن، الكويت، لبنان، فلسطين، سوريا؛

- إفريقيا: الصومال، جيبوتي، جزر القمر، مصر، السودان، ليبيا، تونس، الجزائر، المغرب، موريتانيا.

كما أنها اللغة التي يرجع إليها أكثر من مليار مسلم عبر العالم. اللغة العربية هي اللغة التي يتكلم بها العرب، وهي لغة سامية (مثل العبرية والآرامية والسريانية)، نشأت في جزيرة العرب ثم انتشرت وتوسعت بفضل الفتوحات الإسلامية.

من الضروري بالنسبة للدراسات اللسانية المعمقة ولتلك الدراسات الخاصة بالمعالجة الآلية، بلورة نحو صوري للغة العربية. غير أن هناك عدة مشاكل تصادف الباحثين في هذه العملية منها قلة الموارد والمدونات والذخائر اللغوية المخزنة في شكلها الرقمي والتي يعتمدون عليها في دراستهم، وكذا الإطار الصوري الذي ينبغي اتباعه لإنجاز تلك العملية.

إن الدراسة اللسانية الحاسوبية تعالج اللغة حسب تصورات ومناهج خاصة ليست بالضرورة مختلفة عن التصورات والمناهج اللسانية التقليدية، ولكن أهداف تلك الدراسة هي التي تعطيها طابعا مميزا؛ حيث إنها تحاول إقامة أوصاف صورية صارمة لمختلف الظواهر اللغوية من أجل تزويد الآلة بشتى المعارف والعمليات الموجودة في اللغة.

في أول مؤتمر دولي يقام بشأنها سنة 1965، عُرِّفت اللسانيات الحاسوبية بأنها: «علم جديد تتقاطع فيه اللسانيات مع جهاز صوري تفرزه العلوم المنطقية الرياضية ويخضع للقيود التي تفرضها الآلات المعدة للمعالجة الآلية للمعلومة، ويؤدي البحث في هذا المجال إلى إنشاء نموذج خوارزمي». لقد حددت في هذا التعريف ملامح اللسانيات الحاسوبية التي تتمثل في تلاقي جهازين مفهوميين حديثين نوعا ما هما: اللسانيات في تطورها المتزايد، والعلوم المنطقية الرياضية في رؤيتها الصورية، بمجال

تقني حديث النشأة كذلك هو المعالجة الآلية للمعلومة. كما أعلن هذا التعريف عن الهدف من ظهورها وهو صياغة نموذج خوارزمي وهو خطاطة منطقية رياضية للغة قابلة لأن تشغل في الحاسوب.

رغم ذلك، يُستنتج مما تقدم أن اللسانيات الحاسوبية هي مجال تتداخل فيه التصورات اللسانية والحاسوبية وتتلاقح لتشكل نظريات تعمل على معالجة الوقائع اللغوية وفق منهج حاسوبي لتتمخض عن ذلك تطبيقات متعددة تشمل تلك الوقائع اللغوية لكن في إطارها الآلي. وبالتالي، وإن كانت اللسانيات علما متجذرا في الفكر الإنساني، غير أن ارتباطها بالحاسوب هو من ابتداع القرن العشرين عصر ثورة المعلومات.

مما سبق، يبدو أن الباحثين العرب قد ترجموا هذا المصطلح من المصطلح الإنجليزي ***computational linguistics***، الذي يشير إلى ذلك المجال الذي يتفرع من اللسانيات وعلوم الحاسوب، ويشمل التفاعلات بين اللغة الإنسانية والحواسيب، وهو يتضمن تحليل النص المكتوب والخطاب الشفوي، ترجمة نص أو منطوق من لغة لأخرى، استعمال اللغات الإنسانية (لا الحاسوبية) للتواصل بين الحواسيب ومستخدميها، كما أنه يتضمن نمذجة النظريات اللسانية واختبارها.

وبالتالي، يتضمن موضوع اللسانيات المعلومية تلك الأوصاف اللسانية التي تتسم بالنسقية حيث تشكل عناصرها من أصوات وكلمات وجمل وغيرها وما يقابلها من مدلولات نظامية واطرادا في ظهورها كوقائع لغوية، كما تتسم بالصورية حيث تُستخرج من تلك الوقائع اللغوية المدروسة أشكالها وأمثلتها وتعميماتها وتجريداتها بغض النظر عن الاختلافات التي يمكن أن تسمها لدى تحققها نطقا عند المتكلمين أو كتابة عند المؤلفين.

يرجع ذلك التقارب بين النظريات اللسانية والنظريات الإعلامية إلى سعيها لتحقيق أهداف علمية وتكنولوجية، تتمثل أهدافها العلمية في الإسهام في معرفة بنية العقل البشري وطريقة عمله من

زاوية لسانية، أما أهدافها التكنولوجية فتتضمن بناء حاسوب يحاكي قدرات الكائن البشري اللغوية، وتوظيفه في مجالات تطبيقية متعددة.

يبدو لنا مما تقدم أن مصطلح "اللسانيات الحاسوبية" هو مصطلح عام يشمل كل تلك المجالات المتقدمة والتي تتداخل فيها اللسانيات بعلوم الحاسوب حيث يتناول فيها الباحثون موضوعات لغوية باستخدام الحاسوب، مهما اختلفت المناهج والمقدمات والتوجهات العلمية بين هندسية يطغى عليها التوجه الآلي التقني وبين معرفية يشيع فيها إنشاء النماذج الحاسوبية، سواء في جانبها النظري أو من خلال تطبيقاته المتعددة.

إن أهم ميزات اللسانيات الحاسوبية هي النمذجة الحاسوبية حيث تضم هدف النمذجة وهو وصف السيرورات التي يمكن حسابها والتي يتم فيها تركيب اللغة وتحليلها، بالإضافة إلى تحليل المشاكل في مستواها الحاسوبي، والمقصود بالنمذجة صياغة نماذج أي أنظمة عملياتية تحاكي بنيتها العلائقية سيرورة معينة كسيرورة اللغة.

إن المعالجة الآلية للغات هي حقل بحث متعدد الاختصاصات، يساهم فيه لسانيون وحاسوبيون ومناطقة ونفسانيون ووثائقيون ومعجميون ومترجمون، ويندرج في ميدان الذكاء الاصطناعي. لقد أنشأ المهندسون إواليات يمكن أن تلعب الشطرنج، أو تسوق السيارات أفضل من الإنسان أحيانا، لكن تألية الأنشطة اللغوية الإنسانية، أي جعلها تتم بطريقة آلية، ما زالت تعاني من عوائق، ولا يمكن تحقيقها 100%. ومع ذلك، هناك بعض المهام المحددة التي تستخدم اللغة في المعالجة الآلية يمكن أن تعطي برمجيات جيدة، لكن في أغلب الأحيان تحتاج ضرورة للتدخل البشري قبل عملية المعالجة (من خلال تبسيط الأوصاف مثلا) وبعدها (بالتصحيح مثلا) من أجل الحصول على نوعية مماثلة لما ينتجه الإنسان.

يتناول هذا المجال دراسة اللغة من خلال معطياتها اللسانية التجريبية، فهذه المعطيات هي ذات أشكال متنوعة: إذ يمكن أن تكون نصا مكتوبا أي سلسلة من الجمل ذات محتوى اتصالي متسق، أو حوارا شفويا أو مكتوبا، أو وحدات لسانية أصغر حجما من النص كالجملة أو الملفوظ أو الكلمة... لكن الباحثين يشيرون إلى كل تلك الأشكال عادة بمصطلح "النص" ويجعلونه موضوع دراستهم.

لتجاوز تلك الصعوبات، اقتُرحت عدة استراتيجيات تتمثل خاصة في نمذجة المعارف اللسانية وغير اللسانية الضرورية، والنمذجة الرياضية التي تقوم على الانتظام المشاهد في النصوص، والنمذجة المعرفية التي تقوم على تقييس السلوكات البشرية.

تحتاج المعالجة الآلية للغات إلى موارد حاسوبية متطورة مثل كل القطاعات التي تعمل على حوسبة مشكلاتها المعقدة. بعض البرمجيات تشتغل في الحاسوب، لكن أغلب البرمجيات الذكية منها تستلزم آلات أكثر تقنية من الحاسوب الشخصي؛ إذ لم يمكن الحصول على آلات ذات قدرة كافية على معالجة أحجام ضخمة من النصوص باستخدام معاجم إلكترونية في وقت معقول إلا حديثا مع تطور تكنولوجيا الحاسوب. كما أن الشاشات ولوحة المفاتيح قد كيفت مع أنماط الكتابة والعرض لتشمل الحروف غير اللاتينية والكتابات غير الألفبائية. بالإضافة إلى ذلك، سمح التطور في ميدان الإلكترونيات أيضا بتحسين الأدوات الصوتية كظهور البطاقات الصوتية لتأليف الكلام، وكذا التسجيل الرقمي للتعرف على المنطوق.

الأدوات الحاسوبية من أجل المعالجة الآلية للغات تستخدم في عدة سياقات ولأهداف مختلفة، من أهمها مجال التعليم من خلال تكوين الباحثين، وتوضيح مكنزمات المعالجة الآلية للغات، والبرهنة على اشتغال النظريات النحوية، وكذلك تستخدم في مجال البحث كدراسة تقنيات جديدة حاسوبية ووصفية، وإقامة أوصاف لسانية متطورة للمعالجة، وإنتاج مختلف التطبيقات اللسانية الحاسوبية.

يصنف بعض الباحثين ميادين البحث في المعالجة الآلية للغة الطبيعية إلى أربعة محاور بحثية هي الهندسة اللغوية، واللسانيات المعلومية، والمعلوميات اللسانية، واللسانيات الحاسوبية.

تصنيف محاور البحث في ميدان المعالجة الآلية إلى أربعة محاور يغلب على كل واحد منها أسلوب في التناول الآلي للظاهرة اللسانية وفي الهدف من هذا التناول، أما الهندسة اللسانية فتقوم على استعمال المعلومات اللغوية لأغراض حاسوبية وتطبيقها في مختلف البرامج المنجزة، وأما اللسانيات المعلومية، فتعتبر مختبرا حاسوبيا لفحص مختلف النظريات اللسانية وتكييفها مع الآلة. والمعلوميات اللسانية مجال لإنشاء البرامج الحاسوبية التي تعمل على تقييس النشاط اللغوي الإنساني. وأخيرا، محور اللسانيات الحاسوبية الذي يضم محوري اللسانيات المعلومية والمعلوميات اللسانية إذ يتبع الباحث فيه أسلوبي التكييف والتقييس ويضيف إليهما أسلوبا آخر هو أسلوب النمذجة في تناوله للجوانب الحاسوبية للغة والتي تتعلق بأنظمة بناء الجملة وعناصرها المتعددة في شتى المستويات.

إن توليد الجمل هو ظاهرة لسانية على اعتبار ما تؤول إليه هذه العملية التي تتفرع عنها مجموعة من الجمل ذات طبيعة لغوية إذ هي مكونة من ألفاظ تحمل دلالات معجمية وصرفية وتركيبية، وبالتالي يمكن دراسة نظامها واقتراح النظريات اللسانية في وصفها. كما يمكن اعتبارها ظاهرة لسانية نفسية من وجهة النظر الحركية والدينامية التي تتم فيها عملية التوليد بحيث يقع الربط بين معالجات متتابعة ومترابطة تتناول الكيانات اللغوية وذلك من المستوى الأكثر تجريدا والأقرب إلى الدلالات إلى المستوى الأقل تجريدا والأقرب إلى الألفاظ، وعليه يقترح الباحثون نماذج يتصورون من خلالها الكيفية التي تم بها إجراء هذه العملية. هذا الذي دفعني إلى تخصيص فصل كامل من البحث أتعرض فيه إلى مفهوم توليد الجمل بين اللسانيات واللسانيات النفسية.

إن النظريات اللسانية المختلفة نادرا ما تقدم نفسها بشكل واضح على أنها نماذج لإنتاج الكلام وتوليد الجمل. قد يرجع السبب في ذلك إلى أن البحث في سيرورات إنتاج الكلام وفهمه من مهام علم

النفس اللساني، وليست من مهام اللسانيات التي ينحصر حقلها في دراسة بنية اللغة وانتظام مكوناتها في مختلف المستويات دون العروج على الميكنزمات التي تؤدي إلى معالجة المعلومة اللغوية إنتاجا وفهما وتحويلا وتخزينا. ومع ذلك، يعتمد الباحثون على بعض تلك النظريات اللسانية كأساس نظري من أجل إنشاء نماذج الإنتاج، خاصة تلك النظريات التي تشارك العلوم المعرفية اهتمامها بالملكات المعرفية عند الإنسان والتي تشكل اللغة أبرزها.

يتميز هذا النموذج بأولية التركيب على سائر مكونات النحو، فيما يخص طبيعة اللغة: حيث إن النحو الكلي هو عبارة عن مجموعة من القواعد التركيبية، أو فيما يخص اشتغالية اللغة إنتاجا وفهما. ففي مجال الإنتاج، تعمل البنى التركيبية على توليد الملفوظات أين يتم اختيار الكلمات داخل التركيب الذي هو المكون التوليدي الوحيد، في حين أنه لا مكان لنظرية أو مكون دلالي في هذا النموذج، لأن الدلالة هي نتيجة تأويلية تتبع من إدماج المكونات المعجمية داخل البنية التركيبية. تمثل نظرية العمل والربط آخر ما وصلت إليه النظرية التشومسكية من تطور؛ حيث تهدف إلى إنشاء نحو كلي أي ترمي إلى إنشاء نظرية للخصائص الكلية للغات الطبيعية، وهي نظرية للكفاية اللغوية تأخذ بعين الاعتبار ظواهر اكتساب اللغة.

مختلف النظريات التي تتبع التيار الحركي لها تصور مناقض حول طبيعة المكونات اللسانية ودورها. النحو بالخصوص ليس نظاما صوريا خالصا، ولكنه ترتيب خاص لتشفير المعلومة التحادثية (هذه النظرة الوظيفية)، ويحمل هو نفسه معنى باستمرار مع الدلالة المعجمية (هذه نظرة الأنحاء المعرفية). ترتكز هذه المقاربة على تصور خاص للكفاية اللغوية ووظائفها، وتوحِّد هاتين المدرستين اللتين تحاولان سوية إنشاء قواعد جديدة لعلم نفس اللغة.

بالنسبة للوظيفيين، من الضروري دراسة اللغة في شروطها الطبيعية، دون تجاوز الطبيعة الحقيقية للمكنزمات التي تهدف إلى تمثيل المعارف وإيصال هذه المعارف الممثلة، مما يفترض

الانطلاق من الملفوظات الحقيقية المنتجة في سياق الاتصال، وليس من الملفوظات المولدة مجردة عن السياق. هذه المنهجية نابعة من تصور للنحو بأنه نمط خاص لتشفير اللغة مرتبط باستعمالها، ويسمح بالمعالجة في أفضل الحالات. النحو يشكل نظاما لمعالجة المعلومة متكيف بوجه خاص مع قيود الاتصال اللساني، ويمثل –بوجه محتمل– آخر ما توصلت إليه سيرورة التطور لبناء نظام اتصال إنساني.

تقوم الأنحاء المعرفية على المبادئ المنهجية السابقة نفسها، لكنها تمنح دورا معرفيا بنائيا للاستعمال في إطار نموذج حركي لاشتغالية اللغة. حسب هذا التصور، المفاهيم: الكفاية/الأداء، النظام اللساني (التجريد)/ الملفوظات الخاصة (التجربة) لا يمكن الفصل بينها، والدليل على ذلك أمران، الأول: أن المتكلم ينشئ نظامه اللساني انطلاقا من التجربة المكررة للإنتاج والفهم، من خلال سيرورة التجريد؛ إذ النشاط اللغوي هو نشاط خلاق وحركي من أجل بناء مفهمات، يتم في هذا النشاط إعادة بناء قيمة الوحدات اللسانية وتسويتها دوريا.

الأمر الثاني: أن المكونات اللسانية المعجمية والنحوية لا تشكل بنيات صورية أو مضامين مرجعية بسيطة، ولكنها هي نفسها مفهمات؛ فهي تمثل في الوقت نفسه سيرورات الترميز أي ميكنزمات بناء التمثيلات، ونتيجة هذه السيرورات التي هي نفس التمثيلات في آن معا. وبالتالي، لا يمكن التمييز بين مستوى بنوي ومستوى للمعالجة: كل المكونات اللسانية –بما فيها النحوي– تشكل عمليات لمعالجة التمثيلات. هذه العمليات التي تكيفها التجربة تكشف عن مكنزمات عامة تتدخل في أنشطة معرفية أخرى.

يعد إنتاج الكلام في التصور اللساني النفسي حركة معاكسة لفهم الكلام، حيث تتم هذه العملية بتحويل أفكار المتكلم إلى عبارات. يكاد علماء اللسانيات النفسية يجمعون على وجود ثلاث مراحل أساسية تتم فيها عملية إنتاج الكلام، وهي:

- مرحلة مفهمة الرسالة التي يراد التعبير عنها: وتتحدد فيها المفاهيم والأفكار المطلوب التعبير عنها لفظا حيث يتعلق الأمر بلغة الفكر التي هي مستقلة عن اللغة التي يعبر بها المتكلم.
- مرحلة معجمة الرسالة: حيث يجري اختيار الكلمات الملائمة للمفاهيم وللأفكار التي يرغب المتكلم في التعبير عنها من أجل بلوغ بنى دلالية وتركيبية وفونولوجية وصرفية متماسكة. تسمى هذه المرحلة بالدخول المعجمي الذي يتم في مرحلتين فرعيتين:
  - مرحلة التشفير الدلالي والتركيبي: تتمثل في استرجاع اللمات، وتليها مرحلة التشفير الصرفي الفونولوجي التي تتمثل في استرجاع اللوسيمات.
- مرحلة مفصلة الرسالة: وهي المرحلة الأخيرة التي تقود إلى النطق بها.

إن النماذج الحاسوبية لإنتاج الكلام هي بناءات معقدة ومتنوعة؛ فهي تقدم المعلومات والسيرورات المرتكزة على فرضيات شتى. بالإضافة إلى ذلك، لا تتبنى تلك النماذج الأهداف نفسها؛ إذ لا تتفق فيما بينها على نوع البيانات (مثلا: محاكاة نماذج الأخطاء أو زمن الإجابة)، وعلى الطريقة (محاكاة بسيطة لنتائج عامة أو توزيعات كاملة أنتجها الفاعلون).

يعتبر التوليد الآلي للجمل أحد الميادين المهمة للمعالجة الآلية للغات الطبيعية، نظرا لارتباطه بتطبيقات عدة ولحاجة جل التطبيقات إليه كالترجمة الآلية، وقواعد البيانات. أحاول أن أتعرض في هذا الفصل إلى توصيفه وبيان المشاكل والصعوبات التي يتضمنها، وأن أذكر أبرز المقاربات والمناهج المستخدمة في هذا الميدان.

التوليد الآلي للجمل أو النصوص هو أحد المجالات التي تتناولها المعالجة الآلية للغات الطبيعية وفرع اللسانيات الحاسوبية، حيث يتم فيه توليد صورة أو شكل لساني انطلاقا من معطيات غير لسانية أي انطلاقا من تمثيل صوري لمحتوى دلالي. وبالتالي، يتضمن ميدان التوليد تطوير أنظمة قادرة على

إنتاج نصوص أو ملفوظات متسقة، سليمة ومفهومة. الأنظمة الحاسوبية المنجزة من أجل توليد النصوص بطرق آلية ترجع إلى سنوات 1970.

إن أنظمة التوليد الأولى لم تكن تطمح إلى توليد النصوص ضمن عملية الاتصال، بل كانت تقوم بتفريع جمل ذات محتوى إعلامي عشوائي لأن الهدف كان اختبار نظريات تركيبية، ثم ظهرت بعد ذلك أنظمة لا تستطيع أن تولد إلا جملا منعزلة، وعليه لم يتم تناول مشكل توليد النصوص ضمن عملية الاتصال إلا في حدود 1980.

المولدات هي أنظمة ترتبط غالبا بميدان وتطبيق معينين؛ فالنظام هو « مجموعة متناسقة من الممارسات أو التصرفات، تُنجز بهدف الحصول على نتيجة »، وبالتالي فإن أنظمة التوليد مكونة من مجموعة من المنظومات (أنظمة جزئية) تقدم نتيجة لاشتغالها نصوصا، ويمكن أيضا لنظام التوليد أن يشكل منظومة لنظام أكبر (كنظام الحوار بين الإنسان والآلة...). والميدان هو « الذي يحتضن فنا أو علما أو موضوعا أو فكرة »[3]، لذلك هناك ميادين عامة ترتبط بها المولدات: طبية واقتصادية... وأخرى أكثر حصرا كميدان المناخ البحري والبورصة... والتطبيق هو « عملية الاستناد على شيء، وهو في المعلوميات: برنامج مكتوب من أجل استعمال دقيق »[3]، وبالتالي يمكن تعريف مولد المغرابي مثلا –والذي سيأتي الحديث عنه– بأنه نظام إنتاج وصفات للمطبخ باللغة العربية، أو تطبيق لإنتاج النصوص العربية في ميدان وصفات المطبخ.

على عكس ما يمكن توقعه، لقد نال اللسان العربي نصيبه من اعتناء الباحثين في إنشاء مولدات للنصوص العربية حتى قبل ألسنة أخرى أوروبية كالألمانية والإيطالية، وذلك منذ عهد مبكر في سنوات 1980، لذلك سأتعرض لإحداها والتي قامت بإنشائها الباحثة شادية المغرابي.

تعمل الأنظمة المتوفرة في اتجاه خطي حيث تنتقل من التمثيل الدلالي إلى الشكل السطحي للنص، فتتوالى مختلف الوضعيات في المعالجة في اتجاه واحد، ويتم الانتقال من منظومة لأخرى مع

عدم إمكانية العودة إلى الوراء. هذه المعمارية المنظومية الخطية تمنح للنظام قدرة على التطور والصيانة حيث يتسنى للباحثين تدارك النقائص الموجودة في منظومة من المنظومات التي يتكون منها المولد خلال مرحلة من مراحل المعالجة، واقتراح بدائل لمواطن النقص والخلل.

قد سبقت الإشارة إلى أن نظام التوليد يتكون من مؤلفتين استراتيجية وتكتيكية بحيث تتكلف الأولى بتحديد محتوى النصوص، ومهمة دراسة هذه المؤلفة ترجع إلى العلوم المعرفية، أما المؤلفة التكتيكية فمهمتها منوطة باللسانيات الحاسوبية، وهي التي تعمل على صياغة شكل الرسالة إذ تزود النظام بمختلف الضوابط والقيود التي تعطي لمحتوى الرسالة ومعطياتها المفهومية شكلا لسانيا سليم البناء والتركيب ومقبولا في اللغة المدروسة.

تجدر الإشارة إلى أن **عبد ذياب العجيلي** قد يكون من أوائل الباحثين العرب الذين حاولوا تحليل اللغة العربية باستخدام لغة برولوغ، وذلك في كتابه "الحاسوب واللغة العربية" –لكني لم أتمكن مع الأسف من الحصول على هذا المرجع–، وقد تناول فيه معالجة العربية في مستوياتها التحليلية والتركيبية، والنحوية والصرفية، والدلالية والأسلوبية، والهجائية، بالإضافة إلى تعرضه للترجمة الفورية بين العربية وغيرها من اللغات.

تتألف الجملة من المركبات التي تتركب هي بدورها من وحدات لكل منها رتبة محددة. وقد بينت النظرية النحوية العربية بأن لكل وحدة لسانية موضعا داخل الجملة، وحتى لو تغير ترتيب تلك الوحدات أو موقعها داخل الجملة من أجل أهداف تداولية أو تواصلية أو نصية معينة، فإن مواضعها تبقى محفوظة.

إن الترتيب الأصلي للجمل يبقى مدار اختلاف بين اللسانيين؛ ففيما يعتبر **كرينبرك Greenberg** بأن الرتبة الأصلية في الإنجليزية هي 'فاعل فعل مفعول'، يقترح **مككولي McCawley** أن تكون الرتبة الأصلية فيها هي 'فعل فاعل مفعول'. وأما في العربية، فيميل **كرينبرك**

Greenberg إلى أن الرتبة الأصلية فيها هي 'فعل فاعل مفعول'، وتبعه على ذلك **الفاسي الفهري**، غير أن **تشومسكي Chomsky** ينكر وجود لغات من هذا النمط، اللهم إلا أن تكون محولة من الرتبة الأصلية 'فاعل فعل مفعول'.

النحو **G** المولد للجمل العربية هو نحو صوري من الصنف 2 أو نحو لاسياقي حسب تراتبية تشومسكي؛ بما أن إنتاج اللغة يتم بالتحديد عبر رمز غير نهائي عن اليسار وسلسلة غير معدومة عن اليمين، وأن هذه العملية تتم في كل السياقات وليس فقط في هذا السياق، فكل sg يمكن إعادة كتابته على شكل g g1 مثلا في أي سياق ورد.

لغة برولوغ صممت في بدايات سنوات 1970 على يد **كولمرووِر Colmerauer** وفريقه في جامعة مرسيليا، كان الغرض من تصميمها هو برمجة نظام للاتصال بين الإنسان والآلة، وللمعالجة الآلية للغات الطبيعية. ترتكز هذه اللغة على المنطق الصوري خاصة على حساب المحمولات من الترتيب الأول، لذلك كانت تسميتها نابعة من التركيب بين المقطعين الأولين من البرمجة المنطقية أي "**Prolog**" من "PROgrammation LOGique". وقد كتبت لهذه اللغة منذ تصميمها عدة مترجمات ذات قواعد تركيبية مختلفة، من أشهرها **SWI-Prolog** وهي التي أعتمد عليها عند إعداد برامج توليد الجمل في هذا الجانب التطبيقي.

بالانتقال إلى تمثيل الجمل العربية ذات الشكل: 'عامل معمول1 معمول2' باستخدام نحو العبارات المحددة **DCG**، أحاول تطبيق الإجراءات المتبعة في لغة برولوغ على بعض الأمثلة من الجمل العربية:

relation(إن,الجو,حار).

relation(يكرم,المؤمن,ضيفه).

relation(كان, البحر,هادئا).

- إدخال الحالة العامة:

relation(_عامل, _معمول1,_معمول2).

- ثم أطلب من البرنامج إعطاء قيم لتلك المتغيرات من خلال السؤال الآتي:

?- relation(_عامل, _معمول1,_معمول2).

فيكون رده كالآتي:

_عامل = إن,

_معمول1 = الجو,

_معمول2 = حار;

_عامل = يكرم

_معمول1 = المؤمن,

_معمول2 = ضيفه;

_عامل = كان,

_معمول1 = البحر,

_معمول2 = هادئا.

true.

- ثم أطلب من البرنامج ربط تلك العناصر في جمل:

?- relation(_عامل,_معمول1,_معمول2), write(_عامل), write(' '), write(_معمول1),

write(' '), write(_معمول2), nl.

فتكون الإجابة:

إن الجو حار

إن = عامل_,

الجو = معمول1_,

حار = معمول2_;

يكرم المؤمن ضيفه

يكرم = عامل_,

المؤمن = معمول1_,

ضيفه = معمول2_;

كان البحر هادئا

كان = عامل_,

البحر = معمول1_,

هادئا = معمول2_.

في الحقيقة، هذا العمل التطبيقي الذي أقوم به لا يتناول كل عناصر النظام المولد، ولكنه يتناول جزءا منه وهو الذي يتعلق بتكوين الجمل وتأليفها انطلاقا من المعطيات التي تقدم إلى برنامج برولوغ، أما نظام التوليد فهو أكثر تعقيدا إذ ينطلق من التمثيلات الدلالية ليفرع نصوصا، وبالتالي فهو يحتوي على مؤلفتين استراتيجية وتكتيكية كما مر التفصيل فيهما.

4- **الجمل الاسمية المتكونة من مبتدأ وخبر يكون اسما مفردا:** وهي أبسط صور التأليف في اللغة العربية إلى جانب الجملة الفعلية المتكونة من فعل وفاعل. وحسب النموذج الخليلي: [(ع ← م1) ± م2] ± خ تكون كالآتي:

الملاحظ بأن البرنامج قد قام بتكوين كل الجمل الممكنة انطلاقا من البيانات التي قدمت إليه، وبالطريقة التي أرشد إليها من خلال القواعد، وبالتالي ولَّد جملا سليمة التركيب، لكن اللافت للانتباه أن بعض تلك الجمل سليمة التركيب غير مقبول لدى المتكلمين نظرا لعدم صحة الارتباط بين السمات الدلالية بين المبتدأ والخبر:

5- **الجمل الاسمية المبدوءة بحرف ناسخ:** وهي الجمل الاسمية التي دخل عليها حرف ناسخ ينصب المبتدأ ويرفع الخبر، وهي: إنَّ، ليت، لكنّ، لعل، ولم أذكر "أنَّ" لأن لها أحكاما خاصة، بالإضافة إلى الجمل الاسمية المبدوءة بالحروف المشبهات بـــ"ليس"، إن، ما، لا التي ترفع المبتدأ وتنصب الخبر، لكن أقتصر على أمثلة من النوع الأول. وبالتالي، يكون النموذج على هذا الشكل:

6- **الجمل الاسمية التي خبرها جملة فعلية:** الجملة الاسمية التي يكون خبرها جملة هي التي يسميها النحاة بالجملة الكبرى، ويقابلها الجملة الصغرى وهي التي يكون خبرها مفردا، والجملة المخبر بها قد تكون فعلية أو اسمية، والجملة الفعلية المخبر بها تشبه إلى حد كبير الجملة الفعلية مما جعل بعض النحاة يعدونهما شيئا واحدا، إلا أن حالات المطابقة تبين أنهما مختلفان، وسأتعرض إليها فيما يلي.

أتطرق هنا إلى الجملة الفعلية في أبسط صورها وهي المتكونة من فعل وفاعل. وحسب النموذج الخليلي: [(ع ← م1) ± م2] ± خ تكون كالآتي:

إذن، قام البرنامج بتكوين كل الجمل الفعلية الممكنة، وطابق بين الفعل وفاعله في النوع والعدد، وذلك انطلاقا من القواعد التي زود بها، بيد أنه شكّل جملا غير مقبولة أيضا، وذلك لعدم تزويده بقيود دلالية تمنع تكوين مثل هذه الجمل.