

## المحاضرة 5: حوسبة النظام التركيبي

توطئة: يعد النحو نقطة الالتقاء بين اللسانيات والرياضيات، واللسانيات والبرمجيات، باعتبارهما قنطرة الوصل التي تعبر خلالها مسارات الافتراض المتبادل بين علوم اللغة وعلوم الحاسوب، فضلا عن توفير المعطيات اللازمة للتحليل اللغوي الأعمق وهو الفهم الآلي للنصوص اللغوية.

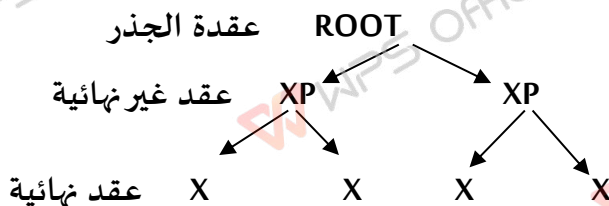
### 1 البنية الذهنية النحوية:

أجمع باحثو اللسانيات العصبية من خلال التجارب أن المنطقة اليسرى من القفص الصدغي الأمامي تطبق نوعا ما من المعالجة النحوية الأساسية، وهذا ما يدعم وجود نمط معين منشأ بالدمغ البشري إلا أنهم اختلفوا في تفسير نوع هذه المعالجة، حيث مازال يكتنفها الغموض.

### 2 التحليل النحوي في إطار المنطق الرياضي:

حاول العلماء تطبيق القوانين الرياضية والمنطقية لنقل هذا النموذج المعرفي الذهني إلى أنساق صورية، لمحاولة تلمس هذا العالم الخفي وإدراك البنية العقلية اللغوية في صورة ملموسة، فانطلقوا جميعا من مسلمة مفادها أن النموذج النحوي الافتراضي يمثل في عدد محدود من العلاقات والقيود، تندرج فيما بينها في هيكل تنظيمي شجري، ثم أثمر تطبيق المنطق الرياضي في صوغه للنموذج النحوي الذهني في تمثيلين نحويين، تعددت في إطارهما النظريات النحوية الحديثة، وهما:

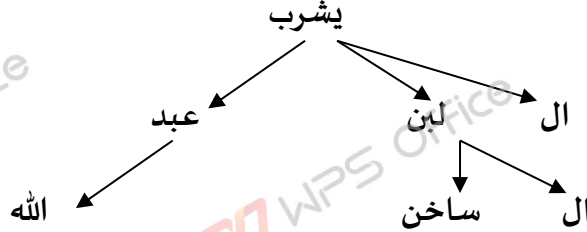
- التمثيل المكوني: وهو بنية شجرية منظمة، تنتظم فيها كلمات الجملة في شكل مركبات أو مكونات نحوية تبعا لنظام نحوي، بحيث تظهر فيها الكلمات كعقد نهائية والمركبات كعقد غير نهائية، وهو ما يعرف بـ "أشجار بنية العبارة"



حيث X هي الكلمات أو الوحدات، بينما XPhrase (XP) تعني المكونات النحوية (مركب اسمي، مركب فعلي،...).

ويتم عادة تنظيم هذه البنية الشجرية أو كتابتها عن طريق التقويم بحيث تظهر العلاقات بينها في صورة اعتمادية، وقد اعتمدها تشومسكي في تمثيله لقواعد النحو، وهي الطريقة المثلى في تمثيل أشجار بنية العبارة حاسوبيا.

- لتمثيل الاعتمادي: هو بنية شجرية منظمة، تنتظم فيها كلمات الجملة مكونة من مسيطر وتابع أو عدة توابع



### 3 أهمية التحليل النحوي:

تكمن أهمية التحليل النحوي للغات العامة -من جانب- في التمكن من التحقق من فرضيات النظريات اللغوية، وذلك وحده عمل تنظيري محكم يبرز طرافة المفاهيم اللغوية والنحوية، أو التأكد من فرضيات قائمة حول لغة معينة، فضلا عن رسم معالم واضحة لنظام الجملة في اللغة المدروسة، وتحديد خصائص علاقاتها التركيبية. واللسانيات الحاسوبية -من جانب آخر-، فهي قوائم تطبيقات معالجة اللغة الطبيعية التي تناظر الأداء الإنساني، والتي يمكن تلخيصها في عنصرين أساسيين:

- المحللات اللغوية: التي تمثل تطبيقات التحليل النحوية (المحللات النحوية) صلب العديد من تطبيقات

معالجة اللغة الطبيعية المختلفة، مثل: الترجمة الآلية، الإجابة على الأسئلة، التدقيق اللغوي، البحث

الدلالي، التخاطب مع قواعد البيانات باللغة الطبيعية، فهم اللغة الطبيعية.

- استخلاص المعلومات: تزايدت المعلومات بصورة متسارعة في ظل تنامي الثورة المعلوماتية وانتشار

الشبكة المعلوماتية التي تتعامل مع العديد من المعارف مثل الكتب، والدوريات، والنشرات،

والصحف، والأبحاث العلمية، وغيرها من دوائر المعارف الأخرى، مما أدى إلى صعوبات جمة في

استخلاصها وتنقيتها من التلوث المعلوماتي، فكان لزاما على المعلوماتيين مواجهة هذه التحديات ببناء

نظم برمجية قادرة على استخلاص المعلومات وتجميعها، تنطلق من المعطيات اللغوية التركيبية

للكشف عن محتوى النص.

#### 4 خطوات التحليل النحوي:

تقتضي أية عملية تحليل لغوي الاستناد إلى عدة طرق إجرائية متدرجة، تنطلق من تحديد النص الخام ثم تقسيمه إلى أجزاء على مستوى الجملة والوحدات، وصولاً إلى تعيين المعلومات الابتدائية المتمثلة في التعرف على أنواع الأجزاء الكلامية التي تساعد بدورها في تعيين العلاقات التركيبية، كما يلي:



• النص الخام/ المدونة اللغوية (corpus): هي كتلة غير منتظمة من النصوص المكتوبة أو المنطوقة،

يمكن التعامل معها آلياً والتحكم في بياناتها ومدخلاتها، ويتم اختيار نصوص المدونة اللغوية المعينة على اعتبارات عديدة يجب مراعاتها في ضوء الهدف المنشود من التحليل، منها:

- انتماء النصوص لأي مستوى (فصيح/ معاصر/ عامي..)

- مصدر النصوص (كتب/ صحف/ مقالات...)

- طريقة اختيار النصوص (الحصر الشامل/ العينات...)

- تنوع النصوص أو التركيز على نوع محدد من النصوص.

• تجزئة النصوص: وتتضمن ما يلي:

- تجزئة على مستوى الجملة

- تجزئة على مستوى الوحدات الرئيسية (كلمة، مركب غير كلامي [عبد الله]/ الرمز أو

العلامة)

- تجزئة على مستوى الوحدات/ العناصر الفرعية (الجدع/ سوابق/ لواحق...)

• العنونة بالأجزاء الكلامية: هي عملية تعيين الأجزاء الكلامية وما تحمله من سمات صرف- نحوية لكل

كلمة منفردة بمعزل عن سياقها الإعرابي في النص، وذلك بإلحاق كل مفردة برموز أو عدة رموز تعبر

عن الجزء الكلامي وما يحتويه من مورفيمات أخرى مثل وقع: VB/ الاختيار: NN.

## 5 موارد التحليل التركيبي للغة العربية وتطبيقاته:

تعد المدونة النحوية للغة العربية التي أنتجتها مؤسسة البيانات اللغوية LDC بنسلفانيا (مدونة بنك بنسلفانيا الشجري) أشهر مدونة للتحليل اللغوي للغة العربية، على الرغم من الفقر الشديد في توفر الموارد اللغوية الموصفة للغة العربية، تليها مدونة براغ الاعتمادي، وهما متشابهتان نسبيا.

ثم قدم مؤخرا مركز أنظمة التعلم الحاسوبي بجامعة كولومبيا بنكا شجريا نحويا ثالثا ينطلق من تقليص حجم المعلومات اللغوية المقدمة في المدونتين السابقتين ومحاولة تجنب المعلومات التي لا فائدة منها بغية تسريع عملية الترميز.