قواعد المعالجة الآليّة للغات الطّبيعيّة:

تعتمد المعالجة الآلية للغات الطبيعيّة بالأساس على مبدأ فهم أنظمة الحاسوب للغة الطبيعيّة وتحليلها؛ فغموض اللغة وعدم وضوحها في البيئات التي تستخدم فيها هذه اللغة، إضافة إلى الإطار العام المحيط بهذه اللغة، يؤدّي إلى عدم فهم الحاسوب للغة البشريّة.

وعندما نطرح مشكلة فهم أنظمة الحاسوب للغات الطبيعيّة فإنّنا نطرح مشكلة الذّكاء الاصطناعي ومدى قدرته على استيعابه لهذه اللغة، ففي بداية عمليّة إخضاع اللغات للمعالجة الآليّة فقد الوسط العام المنشغل بهذا المجال الأمل بإمكانية تمثيل لغة ما بشكل كامل وشامل في برامج الحاسوب وفي مقدرة الحاسوب على فهم الحاسوب للكلمات والجمل مرتبطة بسياقها الاستعمالي، لأنّ الحاسوب يفهم الكلمات والجمل دون ربطها بسياقاتها، وقد كان المشكل المطروح هو محتوى البرنامج المعدّ للمعالجة الآليّة كتمثيل داخلي ولحلّ هذه المشكلة لجأ المختصّون في تصميم هذه البرامج إلى تقسيمها إلى برامج فرعيّة ويحتوي كلّ برنامج على خطوات استدلاليّة.

وتعدّ عمليّة تصميم برامج المعالجة الآليّة للغات الطبيعيّة عمليّة صعبة ومعقّدة جدا؛ لأنّ التحدّي الصّعب الذي تواجهه هذه البرامج هو كيفية دمج فروع مختلفة من المعرفة اللغويّة واستخدامها عمليّا في معالجة اللغة.

وتعتمد برامج المعالجة الآلية مثل البرامج الحاسوبيّة الأخرى بشكل عام على تقنيّات عامّة كتقنيّة النّمذجة وتقنيّة تقسيم البرنامج إلى مكوّنات فرعيّة مستقلّة ومرتبطة في الوقت نفسه عن طريق الخطوات الاستدلاليّة لمواجهة حدّة مشكلة تعقيد النّظام اللغوي، وتقنية الصّورنة حيث يتم صورنة النّموذج اللغوي الاستعمالي وبعد ذلك يتم تصميم النموذج التطبيقي الترميزي الذي يقابل النّموذج الصّوري ويمثّله، وهكذا يصبح البرنامج المبدئي جاهزا لاختبار فعاليته في معالجة اللغة وتسويقه تجاريا لاحقا.

أمظمة المعالجة الآلية للغات الطبيعيّة:

يسعى المختصّون في هذا المجال إلى انتهاج أسهل الطرائق للوصول إلى صيغ رياضيّة لغويّة تمكّنهم من استعمال الحاسوب في معالجة النصوص المكتوبة والمنطوقة والوصول إلى تحقيق هذا الهدف يطرح صعوبتان:

- صعوبة التحليل اللغوي واستعصاؤه.

- وصعوبة إيجاد الأنماط البرمجيّة لهذا التحليل.

وهاتان الصّعوبتان تحتّما ن على مصمّمي برامج المعالجة الآلية للغات الطّبيعيّة التزوّد بنوعين من المعرفة قد يبدو للوهلة الأولى لا علاقة رابطة بينهما وهما: المعرفة اللغويّة الكافية والمعرفة الحاسوبيّة البرمجيّة، أو على الأقلّ التعاون في تصميم البرامج بين اللسانيين ومهندسي البرامج الحاسوبيّة.

وعموما فإنّه للمعالجة الآلية للغات الطّبيعيّة قوانينها وقواعدها وتتكوّن برامج المعالجة من ثلاثة مكوّنات (أجزاء) أساسيّة هي:

أ ـ مكوّن التحليل والتّوليد:

لكي يكون البرنامج ناجحا يجب أن تكون له القدرة على تحليل العديد من العبارات والصّيغ اللّغويّة التي تعبّر عن المعنى نفسه أو المعلومة ذاتها، وفي المقابل توليد العبارات من البيانات المتشابهة.

ب ـ ـ مكوّن تحليل النّصوص والجمل:

والمعلوم انّ الحاسوب يستطيع فهم وتحليل الجمل بمعزل عن سياقها، أمّا فهمها ضمن سياقها فإنّ هذا الأمر يمثّل مشكلا فعليّا امام العقل الآلي.

ج ـ مكوّن التحليل اللغوي:

بعد إدخال نصّ بلغة طبيعيّة ما إلى البرنامج الحاسوبي يتمّ استعمال اللغة المصطنعة لفهم جمل وعبارات النّص؛ حيث تُترجم باستعمال التطبيق الترميزي، ويتمّ تحليل هذه النّصوص المنطوقة أو المكتوبة عبر مراحل متتالية.

مراحل التّحليل اللغوي:

تمزّ معالجة النصوص اللغويّة آليا بالمراحل التالية:

أـ مرحلة التحليل الصّرفي:

ويتخذ هذا التحليل الكلمة مدخلا له؛ حيث يتم تحليل بنية الكلمة لمعرفة جذرها ووزنها وما طرأ عليها من تغيير أو زيادة أو نقصان، وما لحقها من ضمائر أو لواحق، ومعرفة نوع الكلمة نفسها وغير ذلك من الأمور المرتبطة بالجانب الصرفي للكلمة

ب ـ مرحلة التحليل النّحوي: (التركيبي)

وهو التحليل الذي يُعنى بهيكلة الجمل وعلاقتها بعضها ببعض اعتمادا على المعلومات والمعارف النّحوية الخاصّة بلغة طبيعيّة ما، ويرتكز على التحليل الصرفي.

ج ـ مرحلة التحليل الدّلالي:

ويرتكز هذا التحليل على التحليلين السّابقين بهدف تحديد العلاقات الدّلاليّة التي تربط بين أجزاء الكلام وتحديد معنى الجمل والمقصود منها.