

**معالجة المستوى التركيب-دلالي للغة  
بين المقاربة الإحصائية والمقاربة المبنية على نظام القواعد**

**Syntax-semantic level treatment of  
language between the statistical approach  
and the grammar-based approach**

**أ. سالم الرامي**

**الكلية متعددة التخصصات بأسفي  
جامعة القاضي عياض  
المغرب**

[salim.rami@gmail.com](mailto:salim.rami@gmail.com)



## معالجة المستوى التركيب – دلالي للغة بين المقاربة الإحصائية والمقاربة المبنية على نظام القواعد

أ. سالم الرامي

### الملخص:

تبدو المقاربة الإحصائية للغة في تناقض تام مع المبدأ القائم على ضرورة فهم النصّ من خلال تمثيله تركيبياً ودلاليًا بتصريح كامل للقواعد بمثابة شرط أساسي لتحقيق الترجمة الآلية. فلماذا حققت المقاربة الإحصائية فعالية مقبولة بناء على نصوص غير خاضعة لانتقاء مسبق مقارنة بمقاربة القواعد التي يتمّ الاشتغال على تعييدها بواسطة عدد كبير من اللسانيين لمدة سنوات عديدة؟  
كلمات مفاتيح: اللسانيات الحاسوبية- الترجمة الآلية- المقاربة الإحصائية- المقاربة المبنية على القواعد- اكتساب اللغة- العقلانية- التجريبانية- أنموذج اللغة.

### Abstract:

At first glance, it seems that the statistical approach to natural language processing is in complete contradiction to the rule-based systems approach.

Namely, that the latter is based on an explicit declaration of syntactic and semantic rules as a necessary preliminary phase for machine translation.

Hence the legitimacy of asking the following question: why the statistical approach is by far more effective even for non-selected texts in comparison with the results from the rule-based approach designed by a multitude of linguist experts in several years?

**Keywords:** computational linguistic-machine translation-statistical approach-rules-based approach-language acquisition-rationalism-empiricism-language model.

**1- مقدمة:**

نتوخى خلال هذه الورقة المقارنة بين الأنموذج الإحصائي لمستوى التركيب-دلالية والآنموذج المبني على نظام القواعد لذات المستوى. وذلك، في ضوء الترجمة الآلية بمثابة تطبيق لكونها المجال الذي استفاد من الاستثمار الأوفر على مستوى البعدين الزمني والمادي مقارنة بالتطبيقات الحاسوبية الأخرى في مجال اللسانيات الحاسوبية.

**2- حصيلة الترجمة الآلية المبنية على أساس القواعد التصريحية التركيب-دلالية:**

تعتبر النماذج المقدمة في الترجمة الآلية، عبارة عن مزيج بين ألفاظ اللغة الهدف وبين نحو اللغة الهدف المعادل لنحو اللغة المصدر، لتنتهي بالخطأ على كلا المستويين أي الخطأ على المستوى المعجمي بانتقاء لفظ غير مناسب والخطأ على المستوى التركيب-دلالي بترتيب الألفاظ بشكل خاطئ. وذلك، تنزيلاً للمقاربة العقلانية التي تنطلق من الافتراض القائم على أنه يمكن تفكيك اللغة المصدر لعزل أهم وحدات المعنى وترميزها ليتم فيما بعد إعادة تركيبها في شكل لغة طبيعية لتمثيل اللغة الهدف. فيدخل اللسانيون إلى الحاسوب لوائح المفردات وعناصر النحو وعددا كبيرا من قواعد الوصف التحليلية التركيبية. علماً أن هذه القواعد لا يمكن أن تتضمن الاستثناءات وجميع القيود اللازمة لرفع اللبس وتديير الوضعيات الغامضة أو غير المتنبأ بها.

لقد كان فشل الترجمة المبني على ترجمة الألفاظ يبين بوضوح ضرورة سيادة السياق اللغوي قبل كل شيء آخر. إذ تتطلب الترجمة الصحيحة تحديد معنى الألفاظ والتعبيرات المعقدة وفحصاً دقيقاً لتجاوز الألفاظ، وهو ما يتم بواسطة تحليل طبيعي وواضح وغير واع من طرف الإنسان، وهذا أمر في منتهى التعقيد بالنسبة لجهاز آلي بسبب تعدد إمكانات توافق المعاني المحتمل تأليفها وبسبب ضرورة تفعيل مجموعة من المعارف المتعلقة بالعالم الخارجي أي المعارف الخارجة عن السياق اللغوي للجمل، والتي يفتقر إليها الجهاز الآلي. يأخذ ما يسمى بنحو البناء<sup>1</sup> هذا التعقيد بعين الاعتبار بشكل جزئي من خلال ربط المعاني بعناصر مركبة (المركب وما إلى ذلك). وبالتالي، فإن معنى القولة لا يتشكل من تأليف معاني الألفاظ المكونة لها. توقفت الأبحاث المتعلقة بالترجمة خلال التسعينيات وأصبحت أفضل أنظمة الترجمة الآلية هي تلك التي راكمت أكبر عدد من المعطيات اللغوية التي تمكنها من اختيار الترجمة الصحيحة حسب السياق (على سبيل المثال سستران<sup>2</sup>، وهي شركة تأسست في عام 1960).

1- Goldberg, A., *Constructions: A Construction Grammar Approach to Argument Structure*, Chicago, University of Chicago Press, 1995.

2- تأسست الشركة بولاية كاليفورنيا للعمل على ترجمة النص من الروسية إلى الإنجليزية لفائدة القوات الجوية الأمريكية خلال الحرب الباردة. تمت ترجمة أعداد كبيرة من الوثائق وكانت جودة الترجمات كافية لفهم المحتوى لغايات استخباراتية. استمرت الشركة في تطوير مجال الترجمة الآلية إلى الآن. تقدم الشركة خدمة الترجمة الآلية في مجال اللغة العامة وعدد من اللغات الخاصة على موقعها:

<https://www.systran.net>

### 3- حيلة الترجمة الآلية المبنية على أساس النموذج الإحصائي للتركيب - دلالة:

مع بداية التسعينيات، لم يعرف فريق بحث من شركة "إ ب م" أي اهتمام لمشاكل الالتباس الدلالي ولا إلى المعرفة المتعلقة بالعالم الخارجي عن المادة اللغوية، ليستخدّم وبكل بساطة متناثرائي اللغة (ذاكرة ترجمات مجسّدة في نسخ مترجمة طبق الأصل من مداولات البرلمان الكندي، باللغتين الإنجليزية والفرنسية تتيح إمكانية النظر في تقابل النص الإنجليزي والنص الفرنسي على مستوى اللفظ وعلى مستوى المركّب وعلى مستوى الجملة في نفس الوقت). كانت فكرة مصممي النظام بسيطة، تمثلت في إيجاد مكافئات لغوية محتملة مكنت من رسم حدود إمكانات المعالجة الآلية المحضّة القائمة على المنهج الإحصائي على أساس مفهوم تقنية "ن غرام"<sup>1</sup> (سلسلة من الكلمات بطول ن).

على الرّغم من أن نموذج "إ ب م"<sup>2</sup> يبدو بسيطاً بالنظر إلى المهمّة الموكولة إليه، فقد حقّق أداء ملحوظاً مقارنة بتلك الأنظمة الأكثر تفصيلاً والتي استفادت من تراكم قواعد وصفية منجزة من قبل فرق لسانيين لسنوات عدّة. والسبب بسيط نسبياً - وإن كان مضاداً للحدس -، ويكمن في أنّ اللغة تتميزّ بعيد إحصائي أساسي. وبالتالي فمن الممكن إيجاد ترجمة كلمات استناداً إلى ذاكرة متون مترجمة مسبقاً، كما يمكن التعرّف على عبارات اصطلاحية، شريطة أن تكون محلية ومتردّدة نسبياً بواسطة تقنية "المحاذاة غير المتماثلة" التي تُقابل كلمة واحدة في اللغة المصدر بـ "ن" كلمة في اللغة الهدف، وعليه يمكن للخوارزم مثلاً أن يستنتج بأن لفظ بطاطس بالعربية تقابله مراراً العبارة *pomme de terre* باللغة الفرنسية. ولذلك، تعتبر تقنية "ن-غرام" نموذجاً إحصائياً بسيطاً لتمثيل مفهوم السّياق اللغوي بإزالة غموض معنى الألفاظ بشكل طبيعيّ عندما يسمح السّياق اللغوي المحليّ بذلك<sup>3</sup>. أما في الحالات الأخرى (كأن يظهر اللفظ في سياق غير محصّي أو في سياق غير مكتمل التخصيص، ممّا يعني ظهوره على مستوى "ن غرام" غير محدّد سابقاً)، فتتمّ ترجمته حينئذ على أساس المقابل أو المكافئ الأكثر احتمالاً وهو ما يحقّق، مع ذلك، أداءً لا بأس به على العموم. نلاحظ أنّ المنهج المذكور غير قائم على أساس لسانيّ مقنع، غير أنّه يشتغل بنجاح بفضل المتون الرقمية التي تتضمّن الكثير من البنيات الشائعة من جهة، والعامّة أو النمطيّة من جهة أخرى، والتي ستطلّب من اللّساني وصفاً دقيقاً لتقعيد نحو بنائها ووصفاً أكثر دقّة لوصف بعض الحالات منها التي قد تكون غير منتجة في نهاية المطاف. لا نهدف هنا الدخول في تفاصيل فنيّة تتعلّق بتقويم النموذج الإحصائيّ

1- نظرية إحصائية تقوم على أساس تحديد احتمال ورود عنصر ضمن متواليّة من العناصر. فإذا كان حساب الاحتمال يهتم العناصر المجاورة مباشرة لعنصر معيّن تسمّى: غرام واحد وإذا كان حساب الاحتمال يهتمّ العناصر المجاورة على بعد ن من العناصر فتسمّى ن غرام على اعتبار أنّ ن مُتغيّر يمكن أن يأخذ القيمة واحد أو أكثر.

2- يسمّى المنتج اليوم "Watson Language Translator" وهو متواجد بالرباط الآتي: <https://www.ibm.com/fr-fr/cloud/watson-language-translator>

3- Thierry Poibeau, La linguistique est-elle soluble dans la statistique?, Revue Science/Lettres, Épistémologies digitales des sciences humaines et sociales, 2014.

للغة، بقدر ما يهمنّا التسطير على إمكانية تفسير الأساس النظري للأنموذج اللغوي الإحصائي بالنظر إلى فعالية التجريبية في الترجمة الآلية.

لقد شكّل أرشيف الأمم المتحدة المترجم بجودة عالية، المتّين المعتمداً في مقارنة شركة غوغل للترجمة الآلية<sup>1</sup> التي تبنّت بدورها مقارنة شركة "إ ب م". وتجدر الإشارة إلى أنّ المنهجية الإحصائية لغوغل تجد مرجعيتها الفلسفية في قولة لودفيغ فيتغنشتاين (Ludwig Wittgenstein) المعروفة وهي: "ابحثوا عن الاستعمال بدلا عن المعنى". وهو ما يفيد أن الحلول العمليّة هي تلك التي تعطي الأولوية للتقنية على حساب المعرفة مادامت التقنية تمكّن من تشغيل شيء ما دون ضرورة الوعي بالمنهجية المستعملة ودون الأخذ بعين الاعتبار بالعمليّات التفصيلية التي تؤدي إلى النتيجة المطلوبة. ويمثّل لذلك، بكون الطائرة تطير بدون محاكاة منهجية طيران العصافير. وهو ما يميّز في العمق مجال الهندسة التي تكتفي بالجانب النفعي عن المجال الأكاديمي الذي يبتغي وصفا بنيويا لطواهر العالم بشكل نسقي مع تفسيرها لفهمها.

بدون أدنى شك، لا يقوم غوغل بترجمة النصوص بواسطة قواعد تركيب-دلالية وفق المفهوم التقليدي للتركيب وللدلالة. ولكن لا بدّ من التركيز على المرحلة المتعلّقة بتعلّم -أو باكتساب- نظام غوغل للأنموذج الإحصائي للغة، وهو ما يضيّهي أنموذج اللغة خلال مرحلة الاكتساب اللغوي عند الطّفل، لأنّ الاكتساب في كلتا الحالتين المذكورتين عبارة عن تمثّل داخلي لأنموذج لغوي بناء على الاستعمال اللغوي الخارجي سواء كان مجسّدا في متن رقمي خارجي في ما يتعلّق بالأنموذج اللغوي الإحصائي أو كان مجسّدا في الأداء اللغوي للمحيط الخارجي في ما يتعلّق بأنموذج اللغة المكتسب عند الطّفل.

يتضمّن أنموذج غوغل للترجمة قواعد تركيبية ودلالية بشكل ضمنيّ قياسا بضمنية اكتساب الطّفل للقواعد اللغوية من دون الوعي بها. لأنّه في الواقع، يكتسب الأطفال اللغة من دون حفظ قواعد تصريحية للغة ومن دون القيام بتحليل منطقيّ دقيق للجمل، ليتمّ الوعي بطبيعة مقولات الألفاظ بالتمييز بين الأفعال والأسماء في مرحلة لاحقة لمرحلة اكتساب الأنموذج اللغوي. وهكذا، في كلتا الحالتين، يتمّ اكتساب النّحو بالمفهوم اللساني بشكل ضمنيّ وليس بشكل صريح خلال مرحلة التعلّم التي تتجسّد في اكتساب أنموذج إحصائيّ مستمدّ من الممارسة اللغوية المتمثّلة على مستوى المتون بالنسبة لغوغل وخلال مرحلة الاكتساب كما هي معرفة عند التولّيديين في ما يتعلّق بالطّفل. ولذلك، فبطلّب ترجمة العبارة الإنجليزيّة "hot dog" إلى اللغة العربيّة، يتبيّن أنّ النتيجة هي "النقانق" وليست "كلب ساخن". وتجدر الإشارة بوجود التصريح باستنتاج حتميّ مرتبط بطبيعة هذه المقاربة في الترجمة مفاده أنّ ترجمة لغة خاصّة تقتضي اكتساب أنموذج إحصائيّ مستمدّ من متون نفس مجال اللغة الخاصّة، في حين أنّ ترجمة اللغة العامّة تقتضي اكتساب أنموذج إحصائيّ من متون مجال اللغة العامّة. تماما كما هو الحال بالنسبة لمهنة الترجمة حيث نجد مترجمين مختصّين في مجالات محدّدة وآخرين مختصّين في اللغة العامّة. وهو ما يتوافق وتعليل

1- يمكن تجريبه على الرابط الآتي: <https://translate.google.com>

أسباب نجاعة الترجمة الآلية لشركة غوغل الراجع لكون أنموذجها الإحصائي مُستَمَدًا من متن النصوص المترجمة من طرف الأمم المتّحدة التي تتعلّق أساسا بالمجالين الحقوقي والقانوني.

يمكن تعريف الترجمة بكونها انتقاء ألفاظ أو عبارة اللغة الهدف المكافئة للغة المصدر مع ترتيبها بما يتوافق والقواعد التركيبية والدلالية للغة الهدف وهو ما يمكن اعتباره أداء لغويًا بالمفهوم التوليدي يراعي البنية التركيبية والدلالية للغة الهدف. فمثلا ترجمة "avocado salad" من الإسبانية إلى الفرنسية هي "d'avocat salade"، أما في اللغة العربية فهي "سلطة الأفوكادو". ويجد هذا الاختلاف النوعي في الترجمة أسبابه في اختلاف تركيب ودلالة اللغات. بما أن غوغل للترجمة يتعلّم من المتون اللغوية، فجودة الترجمة المُقدّمة سترتفع حتما مع ارتفاع كمّ المتون المُترجمة واختلاف تنوعها، لتبقى مسألة حدود الجودة إشكالية أساسية يستوجب منّا رفعها، فهي بمثابة إشكالية تجريبية جديدة مرتبطة أساسا بالخصائص الصوتية والصرفية والتركيبية للغات. وذلك، لأنّ عدد الألفاظ يعادل عدد المداخل المعجمية بالنسبة للغات الجامدة بينما يرتفع عدد الألفاظ بشكل ملحوظ مقارنة بعدد المداخل المعجمية بالنسبة للغات الإلصاقية والإعرابية، ليصبح كمّ الألفاظ هائلا بالنسبة للغات الاشتقاقية. وبالنظر إلى ما أسلفنا ذكره، سيتضمّن الأنموذج الإحصائي للغات الجامدة اطرادات ترتيب الألفاظ وبالتالي اطرادات ترتيب الألفاظ-داليا، بينما سيتضمّن الأنموذج الإحصائي للغات الإلصاقية بالإضافة للاطرادات المذكورة، اطرادات تطابق الألفاظ صرف-تركيبيا. أما بالنسبة للغات الاشتقاقية فسيتضمّن الأنموذج الإحصائي للغة اطرادات ترتيب الألفاظ التي ستضمّن بالإضافة إلى ما سبق، اطرادات تهمّ قواعد صرف-صواتية. وبطبيعة الحال، كلّما تعدّدت عناصر النظام المجسّدة في الألفاظ وكلّما تعدّدت العلاقات بين العناصر المجسّدة في اطرادات ترتيب الألفاظ، كلّما ستدهور حتما جودة الترجمة الآلية بسبب تعقيد النظام المتضمّن للعناصر والعلاقات بين العناصر.

#### 4- مقارنة تقويمية للمقاربتين المقدّمتين:

بالنظر إلى الترجمة التي يقدّمها غوغل للعبارات الآتية من اللغة العربية إلى اللغة الفرنسية:

Tout d'abord= بادئ الأمر

D'abord= في البداية

C'est réaliste= الأمر الواقعي

la réalité= الواقع

ordre militaire= الأمر العسكري

يمكن من خلال هذه الأمثلة الجزم بأنّ الإحصاء التوزيعي للألفاظ أسلوب ناجح في ترجمة العبارات القصيرة. فكّلما كانت اللغة فقيرة اشتقاقيا كلما كانت الترجمة دقيقة. بينما كانت مقارنة التحليل التركيبي المبني على القواعد التصريحية ستزج بنا في متاهة تحديد الألفاظ المناسبة في اللغة الهدف بناء على دلالاتها وتوافق تلك الدلالات في ما بينها من جهة وتوافق تلك الدلالات مع السياق التداولي من جهة أخرى. فترجمة



عبارة "الأمر الواقعي" يُوَدِّي إلى العبارة "c'est réaliste" في تخطُّ كامل للمستوى التركيبي. ليتبيّن باللموس أنّ المقاربة الإحصائية ناجعة في نمذجة توزيع الألفاظ الجامدة لنستنتج أنّ اللغات الاشتقاقية ستحتاج إلى نماذج إحصائية أضخم، وبالتالي ستحتاج إلى متون تعلم أكبر مقارنة باللغات غير الاشتقاقية. وهو الأمر الذي يقع في التنافي مع الواقع، لأنّ الموارد اللغوية الرقمية المُحَقَّقة متوفّرة ببعض اللغات الأوروبية غير الاشتقاقية في مقابل كمية ضئيلة من الموارد اللغوية الرقمية المُحَقَّقة المتوفّرة باللغات الاشتقاقية بما فيها اللغة العربية. ناهيك عن ندرة المتون في مجالات التخصص باللغة العربية وأحيانا انعدامها باستثناء مجال المعرفة الدينية.

تبيّن لنا أنّ الأنموذج الإحصائي للغة يحتوي ضمناً على تمثيل السياق اللغوي بمفهوم توزيع المحدّد بأنظمة التركيب والدلالة، وهو ما يمنحه قيمة مضافة مقارنة بالأنموذج اللغوي المبني على القواعد التركيبية والدلالية التصريحتين. فهل يعادل السياق التوزيعي للفظ معيّن تمثيله التركيب-دلالي؟ لنعتبر الآن الجملة الآتية: "ما أضّرّ بالعين هو تواجد قطعان الماعز والأغنام بالجوار". يمكن وببساطة ملاحظة أنّه ما من معلومة في السّياق اللغوي للجملة تمكّن من تحديد المعنى المقصود بلفظ العين من ضمن المعاني المختلفة الواردة بالقاموس. ولذلك، لا يمكن توقُّع إمكانية تحقيق ترجمة آلية بجودة عالية على افتراض أنّ الإحصاء التوزيعي للفظ ما يعادل دلالة ذلك اللفظ. ونجد أنفسنا في هذه المرحلة أمام ضرورة توسيع معنى السياق اللغوي باتجاه المادّة اللغوية المتواجدة خارج الجملة لفتح إمكانية صياغة ما يعادل إحصائياً قواعد رفع اللبس الدلالي، في حين أنّه ولحدّ الآن ما من تعريف صوري لمفهوم السياق باستثناء أعمال فبين فونان (*Fabienne Venant*)<sup>1</sup> التي انبنت على اقتراح مفهوم الحساب الديناميكي بمثابة حلّ لإشكالية تعدّد المعنى. تبدو هذه المقاربة في تناقض تامّ مع المبدأ القائم على ضرورة فهم النصّ من خلال تمثيله تركيبياً ودلالياً بتصريح كامل للقواعد بمثابة شرط أساسي لتحقيق الترجمة الآلية. فلماذا حققت المقاربة الإحصائية فعالية مقبولة بناء على نصوص غير خاضعة لانتقاء مُسبق مقارنة بمقاربة القواعد التي يتمّ الاشتغال على تعييدها بواسطة عدد كبير من اللسانيين ولمدّة سنوات؟

بالنظر إلى السؤال المطروح، يُمكن تفرّيعه إلى سؤالين آخرين متداخلين:

أ. هل يمكن تصوّر التركيب والدلالة خارج إطارهما التّقليديّ؟ أخذنا بعين الاعتبار المقاربات الإحصائية التي تمكّنت من معالجة بعض الظواهر اللسانية مثل البحث عن أفضل مقابل للترجمة أو تدبير اللبس المحلّي أو تدبير مسألة تداخل القيود التركيبية المتعدّدة، والتي لم تتمكّن أنظمة القواعد من اقتراح حلول مناسبة لها، تُصبح الإجابة عن السؤال بالإيجاب أمراً مُقنعاً.

1- Fabienne Venant, Le poids des mots, Actes des 7es Journées internationales d'Analyse statistique des Données Textuelles, 2004, France. JADT 04, 2004.

ب. هل يمكن اعتبار الأنظمة المبنية على أساس إحصائي أنظمة خالية من المعرفة التركيبية والدلالية<sup>1</sup>؟ عادة يتم اعتبار الإحصاء والدلالة في وضعية تعارض تام. من جهة هناك الحساب ومن جهة أخرى هناك تمثيل أوصاف لسانية. غير أنّ هذا التقليد القاضي بحكم التعارض ناتج عن حدس خاطئ. في الواقع، يمكن الإحصاء من الأخذ بعين الاعتبار بسياقات الاستعمالات المختلفة للفظ الواحد. كلنا ندرك أنّه من الصّعب تصوّر دلالة لفظ في المطلق بمعزل عن أيّ سياق. وإن افترضنا ذلك فستكون تلك الدلالة بدون معنى<sup>2</sup>، أي وجود تخصيص أو تمثيل دلالي بدون تصوّر عقلي ممكن.

تُقَدِّمُ الدراسات القاموسية الحديثة مقترحات لتمثيل السياقات التركيب-دلالية المختلفة للفظ الواحد، غير أنّ تلك المقترحات تظلّ ذاتية إلى حدّ كبير ولا تعكس استعمالات اللفظ المختلفة المُجسّدة في الواقع اللغوي غير الجامد بحكم طبيعته. من واجهة الأداء اللغويّ أو الإنجاز اللغويّ، كلّ متكلم يستطيع بشكل تلقائي اختيار اللفظ المناسب والمترادفات المتوافقة معه في سياق معيّن. ولذلك، يستوجب تعريف مفهوم السياق من خلال مفهوم الاستعمال اللغويّ للفظ، ممّا يدفعنا إلى طرح أسئلة أخرى:

- كيف يمكننا حصر اطّراد الاستعمال الواحد للفظ معيّن؟

- كيف يمكننا تحديد جميع الاطّرادات المختلفة للفظ معيّن؟

لقد راكمت الصناعة القاموسية أعمالاً مختلفة في هذا الشأن بتخصيص مختلف البنيات التركيبية للفظ الواحد من جهة وبتحديد معايير لتقطيع دلالاته من جهة أخرى، غير أنّ مقترحاتها تتسم بالذاتية بدليل أنّ القواميس لا تتضمن دائما الاستعمالات المختلفة الفعلية للألفاظ.

الظاهر أنّ الإحصاء يقوم بتخصيص البنى التركيبية وتقطيع المعاني أيضا وبشكل آلي وبطريقة جدّ مغايرة وبتناج وثيقة الصلة بمفهوم السياق التركيب-دلالي، لأنّه كلّما كان اللفظ ملتبسا كلّما اتّضح المقصود به بواسطة الألفاظ المجاورة له، سواء تعلّق الأمر بسياق متن اللغة الواحدة أو تعلّق الأمر بسياق تقابل ترجمتين لمتنى لغتين مختلفتين. ففي سياق تقابل متنى اللغة الفرنسية واللغة العربية، فإن العبارة المجازية الفرنسية "faim de loup" يترجمها غوغل إلى اللغة العربية بـ "جائع" بينما تتمّ ترجمة العبارة "un loup faim" بالعربية "ذئب جائع" بينما تتمّ ترجمة العبارتين "faim de lièvre" و "un lièvre faim" بـ "الأرنب الجائع". في حين أنّ ترجمة العبارة المسكوكة الفرنسية "jeter l'éponge" تتمّ ترجمتها بنفس النظام للإنجليزية بـ "to give up" وفي العربية بـ "يستسلم". ممّا يدلّ على أنّ الأنموذج الإحصائيّ للغة يمكن من نمذجة ناجعة ومباشرة لمعالجة ظاهرة المجاز وكذلك ظاهرة العبارات المسكوكة مما يستوجب الجزم معه بأنّ المقاربة الإحصائية لا تقصي التركيب والدلالة كما أنّ حدود جودتها هي نفسها حدود النظريات اللسانية التجريبانية التي تستمدّ وصفها من المتون حصرا، فتكون بذلك غير قادرة على تنبؤ وضعيات لغوية جديدة وهو ما يتناقض والخاصية الإبداعية للغات الطبيعية خلافا للسانيات العقلانية القائمة على نماذج افتراضية وبالتالي تنبؤية.

1-Thierry Poibeau, Traduire sans comprendre? La place de la sémantique en traduction automatique, Dans Langages 2016/1 (N 201), P.77-90.

2- Kayser, D: Une sémantique qui n'a pas de sens. Langages n 87 (septembre 1987) P.33-45.



**المراجع:**

- 1- Goldberg, A., Constructions: A Construction Grammar Approach to Argument Structure, Chicago, University of Chicago Press, 1995.
- 2- Thierry Poibeau, La linguistique est-elle soluble dans la statistique?, Revue Science/Lettres, Épistémologies digitales des sciences humaines et sociales, 2014.
- 3- Fabienne Venant, Le poids des mots, Actes des 7es Journées internationales d'Analyse statistique des Données Textuelles, 2004, France. JADT 04, 2004.
- 4- Thierry Poibeau, Traduire sans comprendre? La place de la sémantique en traduction automatique, Dans Langages 2016/1 (N 201), pages 77 à 90.
- 5- Kayser, D: Une sémantique qui n'a pas de sens. Langages n 87 (septembre 1987) P.33-45.