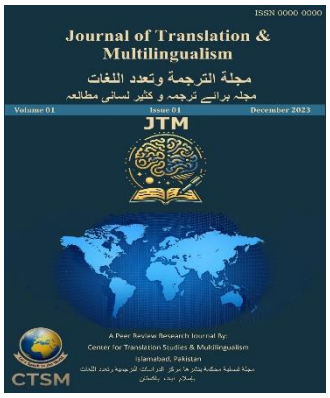


Journal of translation Studies & Multilingualism

مجلة الترجمة وتعدد اللغات

مجلّہ برائے ترجمہ و کثیر لسانی مطالعہ

Homepage: https://ctsmpk.org/index.php/jtm

Volume 01, Issue 01, December 2023

Title:

الترجمة الآلية والتعبيرات المتعددة المفردات:

معالجة ظاهرة الالتباس الدلالي

Machine Translation & Polysemantic phrases:

Tackling the Issues of Ambiguity

Author(s) & Affiliation:

Ali Boulaalam
(Corresponding Author)
FPE- Er-rachidia,
University of Moulay Ismail, Morocco
lingdroit@gmail.com

Azeddine Rhazi,
FLSH-UCA,
Cadi Ayyad University, Morocco
azeddinrhazi@gmail.com

Published on: **15/12/2023**

How to cite: بولعلام، علي وعزالدن غازي " الترجمة الآلية والتعبيرات المتعددة المفردات: معالجة ظاهرة الالتباس الدلالي" مجلة الترجمة وتعدد اللغات، العدد:01، المجلد 01، 15/12/2023 صفحة: 21-36

Publisher: Centre for Translation Studies & Multilingualism, Islamabad, Pakistan

الترجمة الآلية والتعبيرات المتعددة المفردات

معالجة ظاهرة الالتباس الدلالي

Machine Translation & Polysemantic phrases : Tackling the Issues of Ambiguity

Abstract:

The aim of this article is to build the linguistic resources of the Arabic language, according to a methodological conception that combines the different translation approaches (Rul-based, Statistical and Linguistic), in order to make an in-depth description of the Multiword Expressions (MWEs) at the morphological, syntactic, semantic, pragmatic and stylistic levels, as well as the treatment of the ambiguity resulted from to the translation process, as a real parameter to test the operationality of linguistic analyzers describing the Arabic mental representation that could be interacted within the framework of an acculturation translation. The work in progress aimed at the construction of a bilingual dictionary of MWEs using NooJ platform grammar rules and the famous Google Translator.

Keywords : Linguistic Resources ; Semantic Ambiguïty ; MWEs ; NooJ ;Google Machine Translation ; Arabic-French Corpora.

الملخص:

تروم هذه المداخلة الإسهام في بناء موارد لسانية للغة العربية، وفق تصور منهجي يزاوج بين المقاربة اللسانية ومقتضيات الترجمة، وذلك قصد التمكن من إجراء توصيف دقيق للوحدات اللسانية المركبة في مستوياتها الصرفية والتركيبية والدلالية والتداولية والأسلوبية، وكذا البحث عن المحددات الضابطة لظاهرة الالتباس (phenomenon of ambiguity) في علاقتها بإشكالية الترجمة، باعتبارها المقياس الحقيقي لاختبار مدى إجرائية المحللات اللسانية الناظمة لإواياتها، ومدى قدرتها على توفير الأدوات والأساليب المناسبة لتوصيف اللغة العربية داخل نسقية التمثيل الذهني للغات الطبيعية، وجعلها قابلة للتفاعل داخل منظومة المثاقفة في إطار بناء معجم ثنائي اللغة: عربي– فرنسي للتعبيرات المتعددة المفردات.

الكلمات المفتاحية: الترجمة، الموارد اللسانية للغة العربية، الالتباس، التعبيرات المتعددة المفردات،الانزياح الدلالي، معجم ثنائي اللغة: عربي – فرنسي.

1 - السياق العام:

لقد شكلت ظاهرة الالتباس قضية مركزية في الفكر اللغوي قديمه وحديثه، وأثيرت حولها نقاشات همت مختلف دروب المعرفة اللسانية في أبعادها الصرفية والنحوية والدلالية والتداولية والأسلوبية، وقد سعت مختلف الدراسات في هذا المجال إلى تحديد عناصرها وضبط تجلياتها، وذلك من خلال البحث عن الآليات الضابطة لها داخل النسق النظري اللساني، إذ ظل السياق الأداة الرئيسية لمحاولة تجاوز اللبس والغموض الذي يلف المفردة داخل البناء التركيبي، وفق منظومة تدرجية تمتزج داخلها المستويات الصرفية والتركيبية والدلالية. غير أن التعتيم الدلالي الذي يلف التعبيرات المتعددة المفردات (Multi Word Expressions)، باعتبارها بنيات غير تأليفية يجعلها مختلة العلاقة الداخلية التركيبية والدلالية، مما أعاد طرح إشكالية الالتباس خارج دائرة البنية السياقية الداخلية، واستحضر التأويل التداولي الخارجي الذي يستلهم مضامين هذه التعبيرات من خلال الاستعمال، الذي يحصل على الدلالة دون الاعتماد على المفردات المكونة للبنية التركيبية، مركزا في ذلك على البيئة الثقافية المنتجة لها في لغة معطاة وفي لغات طبيعية أخرى يستوجب معها إنجاز دراسات مقارنة لهذه التعبيرات حتى يمكن الاستفادة منها عند الترجمة على أحسن وجه[1].

ومن هذا المنطلق، سنحاول طرح إشكالية الالتباس في علاقتها بدرجة المسكوكية التي تتحدد سلميتها حسب أصناف التعبيرات المتعددة المفردات، وكذا رصد مكوناتها التداولية من خلال عملية التفاعل اللغوي عبر منظومة الترجمة.

2 - دور الالتباس في تحديد درجة مسكوكية التعبيرات المتعددة المفردات:

يقصد بالالتباس احتمال ملفوظ معين لأكثر من تأويل دلالي2، حيث تطرح هذه الظاهرة عائقا أمام عملية الفهم لنص معين سواء خلال عملية القراءة أو التحليل3، وتستوجب ربط مدخل معجمي بأكثر من قيمة وسمية4، وقد أثار هذا المفهوم إشكاليات متعددة تم إدراجها تحت عناوين مختلفة، ارتبطت في أغلبها بظاهرتي المشترك اللفظي والمشترك الدلالي، دأبت الدراسات اللغوية قديمها وحديثها على معالجتهما ضمن ثنائية الحقيقة والمجاز والانزياح

[1]Merella Connena. les expressions figées en français et en Italien : problèmes lexico-syntaxiques de traductions , In contrastes N°10. (1995) p.130.

[2]Spilka, I. Ambiguïté et traduction. *Meta: Journal des traducteurs/Meta: Translators' Journal, 26*(4) (1981).,p. 332.

[3]-Hoceini Youssef et Abbas Moncef : méthodologie multicritère de désambigüisation morphosyntaxique de la langue Arabe, international conférence on arabic language processing, Ecole Mohammadia des ingenieurs, p. 89.

[4]- Ramzi Abbas : la conception et la réalisation d'un concrdancier électronique pour l'arabe. p . 51.

الدلالي بشكل عام، غير أن النظريات اللسانية المعاصرة أثبتت أن الاشتراك الدلالي حالة طبيعية في اللغة وليس انزياحا أو عدولا لغويا5، يحتاج إلى مقاربة علمية في إطار مراجعة هيكلية لبنية المعجم، تعتمد تصنيفا دقيقا للمادة اللغوية حسب الخصائص المورفو- تركيبية للمداخل المعجمية الضابطة لمستوياتها الدلالية رغم تشابهها اللفظي، بيد أن مقاربة الالتباس في علاقتها بالتعبيرات المتعددة المفردات يضاعف - من هذه الزاوية- من حجم هذه الإشكالية المركبة، ويحيل حتما إلى ضرورة التأطير المنهجي- التصنيفي ضمن دائرة "المسكوكية" التي تعد السمة المميزة لهذا النوع من التعبيرات، بوصفها بنيات لغوية صماء توازي نسبة التعابير العادية في جل اللغات الطبيعية، فماذا عن هذه الخاصية ذات الصبغة التركيبية والدلالية والتداولية؟

1-2-مفهوم المسكوكية Idiomaticity :

من خلال رصد جل الدراسات والأبحاث التي تناولت التعبيرات المتعددة المفردات في مختلف اللغات الطبيعية، يبدو أن تحديد خصائصها ومميزاتها يستند إلى ثلاثة مستويات لسانية، يمكن إجمالها فيما يلي:

- **المستوى الدلالي** : إن تأويلها ليس راجعا للكلمات المكونة للجزء المسكوك منها 6، أي عدم ارتباط هذه التعابير بالمعجم في حصر دلالتها المعبر عنها بصورتها المسكوكة7، حيث إن التأويل الدلالي للتعبير لا يرتبط بالوظيفة التأليفية لمكوناته 8، بل تتجاوز فيه الكلمات معانيها الأصلية الدالة عليها في اللغة، وهي منعزلة عن سياقها أو استعمالها في تراكيب للدلالة على معاني أخرى تكتسبها من طريق التركيب أو الاستعمال9، فقولنا – مثلا : رفع عقيرته، بمعنى" شيء جرى اتفاقا، ولا معنى يصل بين الصوت والرجل المعقورة "10.

4 - الحباشة صابر: المشترك الدلالي في اللغة العربية، مقاربة عرفانية معجمية، دار الكتاب الجديد المتحدة، بيروت لبنان، الطبعة الأولى (2015) .،ص:37.

6- M. Cross : syntaxe et localisation de l'information p. 94 – 96.

7- محمد الحناش: ملاحظات حول التعابير المسكوكة في اللغة العربية مجلة التواصل اللساني،المجلد 3، عدد 1، مارس 1991،ص. 31– 32.

8-Nicolas Ruwet : du bon usage des expressions idiomatiques dans l'argumentation en syntaxe générative ,revue québécoise de linguistique vol-13,n°1 p :16.

9 أحمد أبو سعد، معجم التراكيب والعبارات الاصطلاحية العربية القديم منها والمولد، ص: 9.5

10- عبد القاهر الجرجاني، أسرار البلاغة، ص: 245.

- **المستوى المعجمي**: ضعف أو استحالة استبدال أحد عناصر التعبير المتعدد المفردات بعنصر آخر من نفس الفصيلة المعجمية، وذلك في ارتباط بدرجة المسكوكية (idiomaticity)[11].

- **المستوى المورفو- تركيبي**: يعتبر التعبير متعدد المفردات حينما تتعذر استجابته للتغييرات المورفو- تركيبية بشكل مطرد، كما هو الشأن بالنسبة للتعبير العادي حسب الخطاطة الحملية (Predicative Schema) التي ينتمي إليها توزيعا وتحويلا. فثبات الشكل (fixed form) يعتبر من بين سماته المميزة[12]؛ لذلك فانه يصمد أمام التغييرات المورفولوجية من جنس وعدد وزمن واستبدال عائد بآخر، ولا يستجيب كذلك للتحويلات التركيبية بشقيها الصغرى والكبرى من إضمار وحذف وتوسيم وبناء مقلوب و غيرها[13].

وللإشارة، فإن ظاهرة الالتباس داخل الأبنية العادية يتم تدبيرها في نطاق الخطاطات الحملية، باعتبارها آلية ضابطة للعلاقات بين المحمول وموضوعاته، وذلك من خلال ربط المدخل المعجمي بسماته الصرفية والتركيبية والدلالية، أو بعبارة أخرى إخضاعه لعملية الوسم اللساني، غير أن التعبيرات المتعددة المفردات لا تستجيب في مجملها لمبدأ الحملية، نظرا لطبيعة خصائصها اللاتأليفية وصبغتها الدلالية المعتمة، حيث تخضع عملية إنتاجها لكفاية تداولية قائمة على مفهوم الاستعمال، مما حدا بجل الدراسات اللغوية إلى اعتبار هذا النوع من التعبيرات مجرد بنيات لغوية استثنائية، وبالتالي عدم تخصيص حيز لها داخل البنية العامة للمعجم، بشكل يبرز خصائصها وسماتها المميزة. ومن هذا المنطلق، انبرت النظريات اللسانية المعاصرة إلى التعريف بهذه الظاهرة اللغوية، ولاسيما ذات الارتباط بحقل المعالجة الآلية للغات الطبيعية، حيث تبينت صعوبة مقاربتها وفق الأدوات الإجرائية المعتمدة بشأن التعابير العادية ذات الأساس التركيبي التأليفي والمضمون الدلالي الشفاف، ويمكن أن نشير في هذا الصدد إلى بعض النظريات المندرجة ضمن منظومة الأنحاء الترابطية، التي كان لها قصب السبق في وضع الأسس الإجرائية لصوغ الخوارزميات الضابطة للتعبيرات المتعددة المفردات، وذلك من قبيل نظرية المعجم- التركيبي ونظرية معنى- نص.

10- يعتبر شارل بالي من الباحثين الأوائل الذين استعملوا مصطلح "درجة المسكوكية" في كتابه حول الأسلوبية ، حيث ميزوا بين التعابير "المسكوكة كليا" و"المسكوكة جزئيا ،وقد سار في هذا المنحى مجموعة من اللسانيين من ضمنهم غاستون كروس الذي اعتمد نفس التقسيم، مؤكدا أن نسبة "التعابير المسكوكة كليا" جد محدودة في الرصيد اللغوي، وترتبط في أغلبها بالأمثال (في كتابه :التعابير المسكوكة في اللغة الفرنسية، الأسماء المركبة وغيرها) .

12 - Laurence Danlos :la morphosyntaxe des expressions figées ; langages n :63 p :54

14 - علي بولعلام: التعابير المسكوكة والترجمة، دراسة معجمية دلالية، بناء معجم آلي ثنائي اللغة : عربي- فرنسي، أطروحة لنيل دكتوراه الدولة في اللسانيات العربية،2017 ص:47.

2-2-التعبيرات المتعددة المفردات في ضوء النظريات اللسانية:

يمكن القول، إن الأساس النظري للسانيات المعاصرة يتوزع بين توجهين أساسيين، يندرجان على التوالي ضمن أنحاء المكونات والأنحاء الترابطية، وسنحاول مقاربة مسألة الالتباس المرتبطة بالتعبيرات المتعددة المفردات من خلال التركيز على الأساس اللساني الترابطي14، الذي أثبت نجاعته في توصيف الموارد اللسانية، حيث أكد رصانته العلمية وقدرته التفاعلية مع مختلف النظريات التي تدخل ضمن الأساس العلمي لأنحاء المكونات[15]، وذلك من خلال اعتماده على منظور يستهدف بناء معاجم شاملة لمختلف الوحدات اللسانية بشقيها البسيط والمركب، وذلك في ارتباط وثيق بمنظومة المعالجة الآلية للغات الطبيعية، وقد أثمرت هذه الأنحاء نتائج إيجابية، تجلت أساسا على مستوى الأعمال العلمية التي همت مختلف اللغات الطبيعية المؤطرة ضمن نظرية المعجم-التركيبي (Lexicon-Grammar) التي أرسى أسسها زليغ هاريس"Z.Harris" وطور آلياتها المنهجية مورس كروس "M.Gross" وفريقه العلمي بالمختبر الآلي للتوثيق اللساني بباريس 7 بفرنسا، وترتكز هذه النظرية على مفهوم الحملية الذي يعود في أساسه للبنية الفعلية، التي تم من خلالها إجراء عملية تصنيفية للتراكيب اللغوية الأساسية، تؤطرها الأشكال الصورية الستة التالية :

1. ف س°؛
2. ف س°س1 ؛
3. ف س° س1 س2 ؛
4. ف س° س1 ح س2 ؛
5. ف س° ح س1؛
6. ف س° ح س1ح س2 .

وهذه الوظيفة الحملية للفعل تختلف حسب طبيعة البناء التركيبي، حيث يؤطر الفعل العادي الجمل العادية، في حين يؤطر الفعل المركب البنيات المسكوكة، ويقصد بالفعل المركب المنطقة المعتمة أو الثابتة في البنية المسكوكة، التي تشكل مصدر الالتباس والغموض، ولتوضيح ذلك نورد الأمثلة التالية:

التعبير المسكوك	بنيته الصورية	منطقته المعتمة	معناه

[15] – يمكن أن نسوق في هذا الشأن نظرية معنى- نص التي تسعى إلى بناء نموذج لساني، باعتباره جهازا منطقيا يؤطر علاقة الانتقال من مستوى دلالي إلى مستوى صوتي عبر تمثيلات تركيبية و مورفولوجية.

[15] - عز الدين غازي – عمر مهديوي – علي بولعلام: مشروع إعداد موارد لسانية للغة العربية من خلال منصة نوج، وقائع المؤتمر الدولي الأول بالكلية متعددة التخصصات بالرشيدية المغرب 29 – 30 -2016، ض:79.

نفر طائره	ف س°	نفر طائر	ارتبك
عض أنامله	ف س°س1	عض أنامل	ندم
أخذ بيده	ف س° ح س1	أخذ بيد	ساعد
قلب فلان سوادا لبياض	ف س° س1 ح س2	قلب سوادا لبياض	غضب
ضربت عليهم الأرض بالأسداد	ف س° ح س1ح س2	ضربت الأرض بالأسداد	ضاقوا
سلمه ذقنه	ف س° س1 س2	سلم ذقن	استسلم

جدول 1: يمثل التعبير المسكوك وبنيته الصورية ومنطقته المعتمة ثم معناه.

إن التعرف على المنطقة المعتمة يشكل المدخل الرئيس لفك الالتباس على مستوى التعبيرات المتعددة المفردات، وتشمل هذه المنطقة غالبا الفعل وأحد مفعولاته سواء المباشرة أو غير المباشرة، ويتحكم في تحديد دلالتها الاستعمال والتداول، بدلا من الدلالات المعجمية للمفردات المكونة لها، هذه الخصوصيات الواشمة لهذا النوع من التعبيرات يقتضي إدماجها داخل المعجم باعتبارها وحدات مستقلة قصد تيسير التعرف على معانيها بيسر.

وعلى ضوء هذه المعطيات، وانسجاما مع الأدوات الإجرائية التي توفرها لسانيات المنصات ذات الأساس اللساني الحاسوبي، فقد تم التمييز بين أربع وحدات لسانية ذرية، يمكن إجمالها فيما يلي:

1) المورفيمات: تشكل وحدات لسانية تظهر في النصوص على شكل سوابق ولواحق ووسائط وجذور، ولا تميزها بالضرورة فواصل بينها(Separators)؛
2) المفردات البسيطة: تعتبر وحدات لسانية ترد في النصوص كمتواليات من الحروف والفواصل؛
3) المفردات المركبة: عبارة عن وحدات لسانية ترد في النصوص على شكل حروف وفواصل؛
4) التعابير غير المتصلة: تتشكل من متواليات غير متصلة ترد في النصوص على شكل حروف وفواصل[16].

وتنحصر التعبيرات المتعددة المفردات في الصنفين الأخيرين، حيث تضم بالإضافة إلى التعابير المسكوكة، المتلازمات والمفردات المركبة والوحدات الاسمية وأسماء الكيانات وبنيات الفعل العماد والبنيات المتعدية بحرف جر والمصطلحات (Pragmaterm)، وتتفاوت درجة مسكوكيتها حسب معيار الاتصال والانفصال، حيث تصطف هذه الأصناف برمتها في مجموعتين رئيستين؛ تندرج الأولى ضمن التعابير المُعْجَمَة (Lexicalised Expressions)، وتنضوي

[16]-Max Silberztein : la formalisation des langues, l'approche de Nooj p :95.

الثانية تحت النمط التعبيري التواردي (Instutionalised Expressions)، وقد تجاوزت بذلك الحدود الضابطة للمفردة في ذاتها وكذا في ارتباطها بغيرها داخل المنظومة التركيبية والدلالية، حيث تتميز بخصائص يطبعها عدم التأليف والغموض الدلالي، مما يجعل منها كتلا لغوية صماء يصعب تحليلها وتوليدها وفق القواعد الناظمة للبنيات اللغوية العادية، كما يضاعف من الإمكانات الإجرائية لترجمتها ومعالجتها آليا.

3 - ظاهرة التباس التعبيرات المتعددة المفردات وإشكالية الترجمة:

إن إجراء عملية ترجمة التعبيرات المتعددة المفردات تطرح صعوبات جمة، بل تعد مؤشرا دالا على التباسها وارتفاع منسوب غموضها، بل تشكل تحديا حقيقيا بشأن تعلم لغة أجنبية أو بالنسبة لأنظمة الترجمة بشقيها العادي والآلي[17]،حيث تتجاوز التقابلات بين المفردات للغتين المصدر والهدف ذات الصبغة الحرفية، التي ترتكز على معاجم المفردات في بنائها، وأصبحت مسألة التقابل (équivalence) مثار نقاش على مستوى الإنتاج الترجمي، وذلك في إطار التفاعل بين اللغات الطبيعية بصيغه الثنائية والمتعددة. حيث لم يعد الأمر يتعلق فقط بتقابل أحادي تحكمه العلاقات التقابلية بين المفردات بين اللغتين المصدر والهدف، بل تجاوز ذلك إلى أصناف متعددة من هذا التقابل ترتبط بطبيعة التعابير المراد ترجمتها، وهذه الأنماط من التقابل يمكن إجمالها فيما يلي[18]:

1. التقابل المباشر: ويتسم بالترجمة الحرفية من خلال البحث عن المفردات والتعابير المماثلة في اللغة الهدف؛
2. التقابل غير المباشر: ويتميز بالبحث عن الخصائص المورفو تركيبية المميزة للغة الهدف خلال عملية الترجمة؛
3. التقابل المسكوك: ويتعلق بالتعبيرات المتعددة المفردات، التي تستوجب معالجة خاصة باعتبارها وحدات لغوية متراصة، تستوجب ترجمتها بشكل كلي من خلال الارتكاز على دلالتها الاستعمالية.

وبناء على ما سلف، فإن التقابل المسكوك يرتكز على الدلالة الاستعمالية، التي لا تتحصل من مجموع دلالات العناصر المكونة للتعبير، وبذلك يتم تحديد معنى التعبير المتعدد المفردات كلية، ثم البحث عن مقابله في اللغة الهدف، عبر ثلاث إمكانيات، يمكن إجمالها فيما يلي:

[17]-Danlos Laurence. Présentation : lexique-grammaire des expressions figées. In: Langages, 23e année, n°90, 1988. p. 5.

[18]Sułkowska, Monika. (2008). Expressions figées dans une perspective multilingue: problèmes d'équivalence et de traduction. *Neophilologica*,p : 198.

1) ترجمة تعبير مسكوك من خلال البحث عن مقابل مسكوك في اللغة الهدف، ويمكن أن نستدل على ذلك من خلال أمثلة في سياق إجراء عملية ترجمة التعبيرات المسكوكة العربية إلى اللغة الفرنسية:

التعبير المسكوك العربي	المقابل المسكوك الفرنسي
وضع حمله	Débarrasser(se) de ses lourdes charges
يسبح ضد التيار	nager contre le courant
يفتت الأكباد	Causer de la douleur
وضع اللمسات الأخيرة	Mettre la main finale

جدول 2: يمثل ترجمة التعبير المسكوك من العربية إلى الفرنسية من خلال البحث عن مقابل مسكوك

ونلاحظ من خلال هذه الأمثلة، أن إجراء عملية ترجمة التعبيرات المسكوكة بمقابلات مسكوكة في اللغة الهدف، تتضمن في بعض الأحيان صنفا من المقابلات تتكون من وحدات معجمية مماثلة، وتحمل دلالة مسكوكة، من قبيل: يسبح ضد التيار، nager contre le courant، وهذا يندرج ضمن عملية المثاقفة والتفاعل بين اللغات الطبيعية، التي تشكل الترجمة إحدى قنواتها الرئيسية، كما يؤكد في حالة عدم وجود مقابلات مسكوكة في اللغة الهدف تدل على الدلالة التداولية المطلوبة على الاختلاف والتنوع الثقافي.

2) ترجمة تعبير مسكوك بمفردة في اللغة الهدف، من قبيل:

التعبير المسكوك العربي	المفردة المقابلة بالفرنسية
لقي حتفه	mourir
أفل نجمه	échouer
تهلل وجهه	réjouir

جدول 3: يمثل ترجمة تعبير مسكوك بمفردة في اللغة الهدف.

3) ترجمة تعبير مسكوك بتعبير عادي في اللغة الهدف، كما في الجدول التالي:

التعبير المسكوك العربي	التعبير العادي المقابل باللغة الفرنسية

Se mettre en colère	انتفخت أوداجه
obéir par nécessité	جاء مكره أخوك لا بطل
Vivre dans une situation difficile	عاش بين قوسين

جدول 3: يمثل ترجمة تعبير مسكوك بتعبير عادي في اللغة الهدف.

يتبين من خلال ما سبق، أن ترجمة التعبيرات المتعددة المفردات تقتضي حتما بناء قواعد معطيات شاملة تتضمن مقابلات ثنائية أو متعددة على مستوى اللغة الهدف، حسب الإمكانات المطروحة وفق الترتيب المحدد أعلاه، هذه العملية ستمكن لا محالة من تجاوز الأعطاب والصعوبات التي تعرفها منظومة الترجمة عامة، والترجمة الآلية خاصة.

4 - المقاربة اللسانية الإحصائية في الترجمة الآلية ومعالجة الالتباس:

تخضع عملية الترجمة بشكل عام لإجراءات محددة، حيث يتم في البداية تقطيع (Segmentation) النص إلى جمل ثم إلى كلمات، من أجل تيسير البحث عنها في المعجم المتضمن في النظام وتحليلها، ثم يقوم البرنامج بتحديد الخصائص النحوية والمقابلات الممكنة في اللغة الهدف، مع تحديد الوظائف التركيبية لمختلف مكونات الجملة كالفعل والفاعل والمفعول وغيرها وإرفاقها بقيمها الدلالية. وفور القيام بهذه الإجراءات يتم الشروع في عملية النقل المعجمي بين اللغتين مع الأخذ بعين الاعتبار السياق، ثم الترتيب المحلي للكلمات وفق مستلزمات النقل التركيبي[19].

وللإشارة، فإن ترجمة التعبيرات المتعددة المفردات آليا، لا يمكن إجراؤها دون صوغ خوارزمات تمكن من التعرف عليها، هذه العملية ترتكز بالضرورة على تحديد مناطقها المعتمة[20]، ثم إخضاعها لعملية تقطيع تتجاوز تقنية البياضات وعلامات الترقيم، باعتبارها آلية شكلية أثبتت جدواها بالنسبة للتعابير العادية، واعتماد تقطيع قائم على مكون دلالي ترابطي (Chunking) يضم أحيانا أكثر من مفردة واحدة، غير أن هذه المسألة تطرح صعوبات بالنسبة للتعبيرات المنفصلة، استطاعت بعض المنصات اللسانية وضع خوارزمات ضابطة لها في إطار منظومة الأنحاء المحلية.

17 - علي بولعلام: التعابير المسكوكة والترجمة، دراسة معجمية دلالية، بناء معجم آلي ثنائي اللغة : عربي- فرنسي، أطروحة لنيل دكتوراه الدولة في اللسانيات العربية، 2017، ص: 287.

20- Laporte Éric. Reconnaissance des expressions figées lors de l'analyse automatique. In: Langages, 23e année, n°90, 1988. p.117.

وقد أثبتت التجربة أن بعض أنظمة الترجمة الآلية تتعثر خلال عملية ترجمة التعبيرات المتعددة المفردات، وجوهر هذه الوضعية يشكك في الأساس النظري التي تقوم عليه، حيث تعتمد جلها على المقاربة الإحصائية لوحدها والتي أثبتت عدم نجاعتها، ويظل بذلك الأساس اللساني المحدد الرئيس إلى جانب التقنية الإحصائية في إنجاح عملية الترجمة الآلية، التي ما تزال تتلمس الطريق نحو تحسين أنظمتها وترقية أدواتها من خلال الارتكاز على تصورات نظرية ومنهجية ذات منحى لساني-إحصائي (conception théorique hybride) تدمج موارد لسانية خارجية كلما تطلب الأمر ذلك، ولتوضيح هذه المنهجية نورد تطبيقاً من خلال بعض أنظمة الترجمة الآلية مثل:

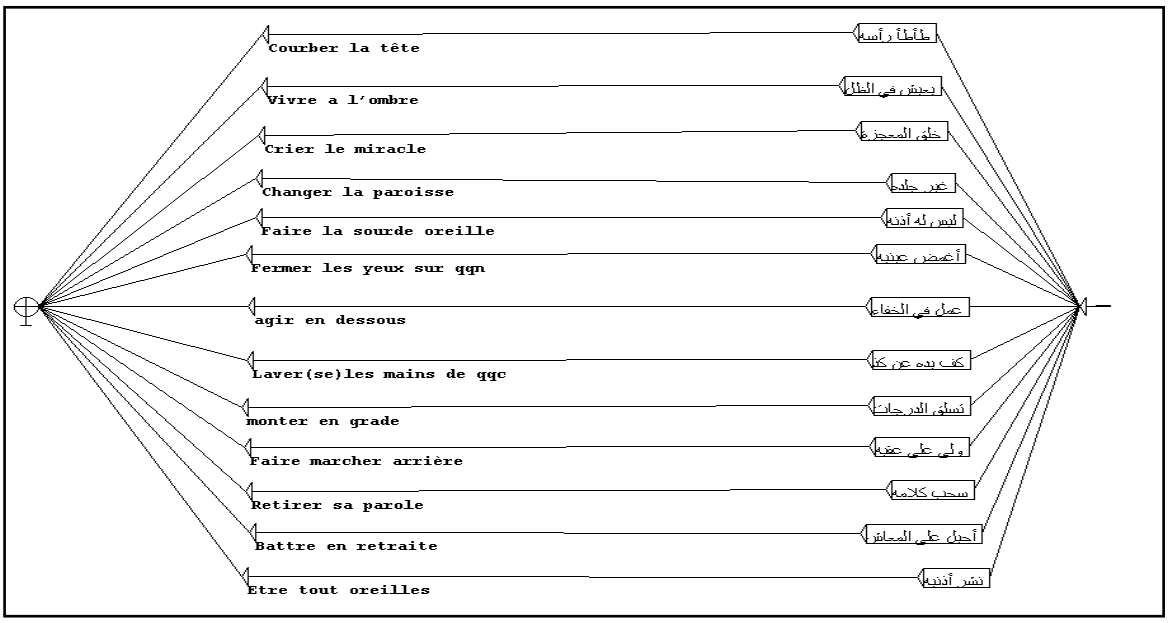
- نظام غوغل"google" للترجمة كما في الشكل التالي :

شكل 1: يمثل ترجمة غوغلgoogle لمجموعة من التعابير المتعددة المفردات.

يتبين من خلال هذا التطبيق، أن نظام غوغل قد أخفق في وضع مقابلات صحيحة لترجمة تعبيرات مسكوكة عربية إلى الفرنسية، سواء عبر مقابلات مسكوكة أو تعابير عادية أو كلمات حاملة لنفس المعنى، مما يستلزم مراجعة دقيقة لبرنامجه بشكل يأخذ الخصائص والسمات المميزة للتعبيرات المتعددة المفردات، التي تمثل نسبة توازي التعابير العادية على مستوى الرصيد اللغوي الطبيعي. وفي هذا السياق، انبثقت موجة من أنظمة الترجمة الآلية في سياق جيل لسانيات المنصات، استطاعت لملمة مختلف الظواهر اللغوية وفق تقنيات المدونات الشاملة والمعاجم الآلية التي تتضمن الوحدات اللسانية بشكلها البسيط والمركب والمتصل والمنفصل، وفق أسس تعتمد تقنيات الوسم

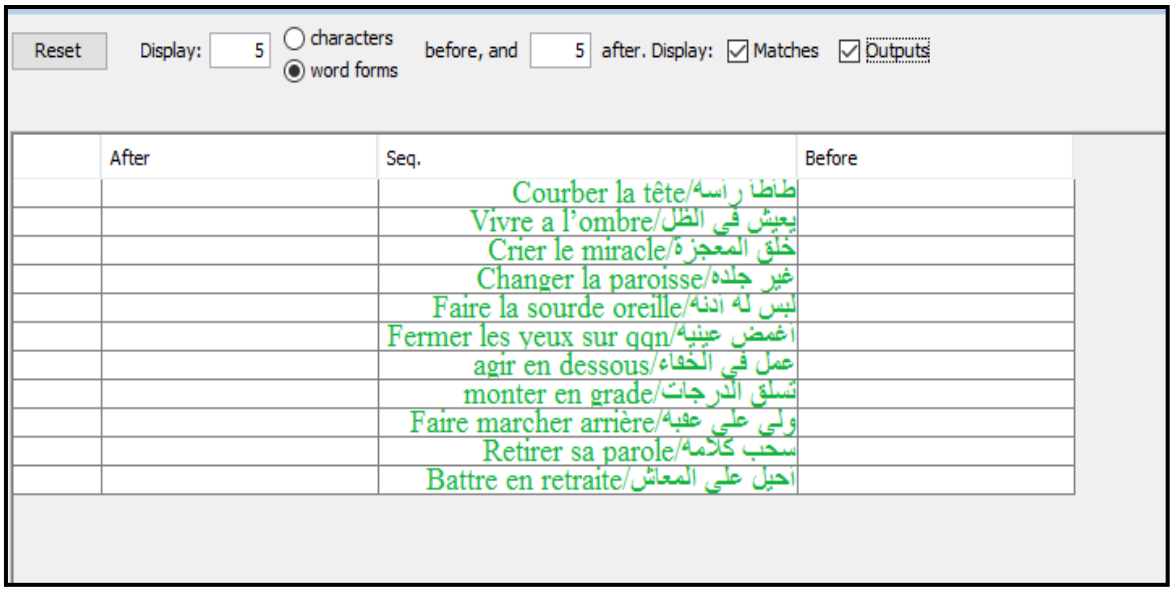
المحددة لمختلف الخصائص والسمات المميزة للوحدات اللسانية الذرية (Unités Linguistiques Atomiques) ويمكن أن نشير في هذا الصدد إلى منصة نوج" NooJ"[21].

- منصة نوج" NooJ ":



شكل 2: تقنية الترجمة بواسطة الرسوم (graphes).

[21] منصة نوج NooJ بيئة تعنى بتطوير اللغات الطبيعية وتطبيقاتها الحاسوبية، مجهزة بأدوات لسانية بمستوياتها المختلفة حيث تساعد :على بناء الموارد اللسانية : تجميعا وتصنيفا وبنينة، و قادرة على صورنة (formalisation) مدونة ضخمة في شتى الظواهر اللغوية: الإملائية والخطية والنحوية والصرفية (إعراباً واشتقاقاً) والمداخل المعجمية بسيطها ومركبها وبتعبيراتها المتعددة المفردات (المتصلة منها والمنفصلة)،
علاوة على قدرتها على صورنة البنيات الصرفية والتحويلات التركيبية بالتوسيمات (annotations)الدلالية التي تستجيب إلى أنحاء "نوج" الالكترونية ، والهدف الرئيس لهذه المنصة هو معالجة اللغات الطبيعية، والتعليم والترجمة الآلية. للمزيد انظر : Silberztein, M. (2016), Formalizing Natural Languages: The NooJ Approach, First Edition. © ISTE Ltd 2016. Published by ISTE Ltd and John Wiley & Sons, Inc.



شكل 3: محصلة عملية الترجمة كما تظهر من خلال الكشاف السياقي لمنصة نوج.

إن انتقال اللسانيات من نماذج (modèles) قائمة على التفسير إلى نمذجة (modélisation) صورية قائمة على توصيف تصنيفي دقيق لبناء موارد لسانية وفق أسس نظرية وإجراءات منهجية، قد مكن من تجاوز جملة من الإكراهات التي شكلت عائقا حقيقيا أمام المعالجة الآلية للغات الطبيعية، وتؤكد هذا المنحى منصة نوج ذات المصدر المفتوح، الذي يكفل إدماج الخاصيات المعجمية-التركيبية لمجموعة من اللغات الطبيعية في نسق صوري موحد.

5 - خلاصة تركيبية

إن بناء نظام للترجمة ثنائية كانت أم متعددة اللغات، يقتضي الاعتماد على مجموعة من الإجراءات المنهجية تتمثل في المنحى الإحصائي- اللساني، وذلك باستغلال الأساس النظري الوصفي للغات الطبيعية المتمثل في نحو المكونات والنحو الترابطي وباعتماد مدونة لسانية شاملة لجميع بيانات التعبيرات المتعددة المفردات بخصائصها الدلالية والمعجمية والمورفو- تركيبية، تمثل الجملة فيها الوحدة الأساسية في عملية الترجمة كما يفعل نظام نوج Nooj، علاوة على اعتبار الأنساق الثقافية بإبعادها التفاعلية المنتجة لهذه البنيات والتي تساعد على فهم سليم وطبيعي للغة ما ونقل محتواها الثقافي والحضاري إلى لغات وثقافات أخرى هدف، مما يضاعف من الإمكانات الإجرائية لترجمتها ومعالجتها آليا.

إن مسألة الالتباس الدلالي الناتج عن ترجمة التعبيرات المتعددة المفردات من العربية إلى الفرنسية - كما لاحظنا ذلك في تجربتنا مع منصة نوج- تحددها الطبيعة الصماء لهذه البنيات والتي تتقلص كلما اعتمدنا المنهجية الإحصائية-اللسانية في بناء الموارد اللسانية ، ووضحنا بنماذج تطبيقة قوة وإنجازية وجودة منصة نوج التطويرية لما لها من أهمية في :

- تطوير التطبيقات لمعالجة اللغة صرفياً وتركيبياً ودلاليا وتحليل المدونات بالنسبة للمختصين في الترجمة الآلية في إطار متعدد المعالجات (multi-agent).

- استغلال أدوات المعالجة الآلية للغة مثل نظام إدارة البيانات ولغة البرمجة مفتوحة المصدر، التي تمنح إمكانية معالجة المدونات المعجمية متصلة أم منفصلة كانت أم متعددة المفردات بطرق البحث والاختبار ودمج الموارد الخارجية كلما تطلب الأمر ذلك، مع إمكانية التحقق من البيانات المستخرجة لضمان جودتها بتصحيحها إذا لزم الأمر، وهذا ما يتطلب صورنة مكثفة للغة.

- تحديد المناطق المعتمة لهذه التعبيرات، ثم إخضاعها لعملية تقطيع تتجاوز تقنية البياضات والعلامات، باعتبارها آلية شكلية أثبتت جدواها بالنسبة للتعابير الحرة، وذلك باعتماد تقطيع قائم على مكون دلالي ترابطي (chunking) يضم تارة أكثر من مفردة، غير أن هذه المسألة تطرح صعوبات بالنسبة للتعبيرات المنفصلة.

- اعتماد نوج لخوارزمات ضابطة في إطار منظومة الأنحاء المحلية تصاغ يدويا وتمثل إلكترونيا بالرسوم (graphes) والمحولات (transducteurs).

وانسجاما مع الأدوات الإجرائية التي توفرها لسانيات المنصات ذات الأساس اللساني الحاسوبي، كما مع مقاربة نوج ، فقد تم التمييز بين أربع وحدات لسانية ذرية متصلة وغير متصلة، وهي: المورفيمات والمفردات البسيطة: والمفردات المركبة: والتعابير غير المتصلة: تتشكل من متواليات غير متصلة ترد في النصوص على شكل حروف وفواصل ويضم الصنفين الأخيرين، بالإضافة إلى التعابير المسكوكة، المتلازمات والمفردات المركبة والوحدات الاسمية وبنيات الفعل العماد والبنيات المتعدية بحرف جر والمصطلحات، وتتفاوت درجة مسكوكيتها حسب معيار الاتصال والانفصال، حيث لاحظنا كذلك أن التعبيرات المتعددة المفردات تندرج ضمن صنفين رئيسيين هما التعابير المعجمة

(Lexicalised Expressions)، والتعابير التواردية (Instutionalised Expressions)، وقد تجاوزت بذلك الحدود الضابطة للمفردة في ذاتها وكذا في ارتباطها بغيرها داخل المنظومة التركيبية والدلالية.

المصادروالمراجع

أ-باللغة العربية:

1. الحباشة صابر: المشترك الدلالي في اللغة العربية، مقاربة عرفانية معجمية، دار الكتاب الجديد المتحدة، بيروت لبنان، الطبعة الأولى 2015.
2. محمد الحناش: ملاحظات حول التعابير المسكوكة في اللغة العربية مجلة التواصل اللساني، المجلد الثالث، العدد الأول، مارس 1991.
3. عبد القاهر بن عبد الرحمن بن محمد, عبد القاهر الجرجاني. (1991). أسرار البلاغة. جدة،السعودية: دار المدني.
4. عزالدين غازي – عمر مهديوي – علي بولعلام، مشروع إعداد موارد لسانية للغة العربية من خلال منصة نوج Nooj، وقائع المؤتمر الدولي الأول بالكلية متعددة التخصصات بالرشيدية المغرب 29 – 30 -2016.
5. علي بولعلام، التعابير المسكوكة والترجمة، دراسة معجمية دلالية، بناء معجم آلي ثنائي اللغة:عربي-فرنسي،أطروحة لنيل دكتوراه الدولة في اللسانيات العربية،2017.
6. غازي عز الدين، بناء المعجم الآلي للغة العربية: معالجة التعبيرات المتلازمة، أطروحة لنيل الدكتوراه، كلية الآداب عين الشق، الدار البيضاء 2009.
7. غازي عز الدين، دور التعابير المسكوكة في الترجمة الآلية، مجلة ترجمان، م 7 ع 1 سنة 1998.
8. غازي عز الدين، رهانات نظرية حاسوب– لسانية في بناء معجم آلي للتعابير المسكوكة. مجلة الدراسات المعجمية، العدد السادس، يناير 2007.

ب- باللغات الأجنبية:

1. Abbes, R. (2004). La conception et la réalisation d'un concordancier électronique pour l'arabe (Doctoral dissertation, Thèse de Doctorat, Institut national des sciences appliquées de Lyon, Lyon, France).
2. Hoceini, Y., & Abbas, M. (2009, May). Méthodologie Multicritère de Désambiguïsation Morphosyntaxique de la langue Arabe. In 3rd International Conference on Arabic Language Processing, CITALA'09.
3. Laporte Éric. Reconnaissance des expressions figées lors de l'analyse automatique. In: Langages, 23e année, n°90, 1988.
4. Laurence Danlos :. Présentation : lexique-grammaire des expressions figées. In: Langages, 23e année, n°90, 1988.

5. Laurence Danlos :la morphosyntaxe des expressions figées ;langages n :63 .
6. Merella Connena. (1995), les expressions figées en français et en Italien : problèmes lexico-syntaxiques de traductions , In contrastes N°10. 95.
7. Nicolas Ruwet : du bon usage des expressions idiomatiques dans l'argumentation en syntaxe générative ,revue québécoise de linguistique vol-13,n°1.
8. Gross, M. (1983). Syntaxe et localisation de l'information. A. Lichnerowicz, F. Perroux & G. Gadoffre (éds), Information et Communication, Paris: Maloin.
9. Silberztein, M. (2015), La formalisation des langues: l'approche NooJ. ISTE 2015 .
10. Spilka, I. (1981). Ambiguïté et traduction. *Meta: Journal des traducteurs/Meta: Translators' Journal*, *26*(4) .
11. Sułkowska, Monika. (2008), Expressions figées dans une perspective multilingue: problèmes d'équivalence et de traduction. *Neophilologica*.
12. Rhazi A and Boulaalam A.(2018) ,Corpus based Extraction and Translation of Arabic MWEs,NooJ Proceedings 2017, Springer CCIS 811.pp.13 .