

Impact de l'intelligence artificielle (la traduction automatique) dans l'apprentissage des langues : cas du Tamazight

GUERCHOUH Lydia^{1*} 

¹Université de Tizi-Ouzou Mouloud Mammeri, Algérie

lydia.guerchouh@ummt.dz

Mehmet Hakki Suçin² 

²Université d'Ankara, Turquie

mhakkisucin@yahoo.com

Reçu: 02/01/2024,

Accepté: 03/03/2024,

Publié: 30/06/2024

Impact of Artificial Intelligence (Automatic Translation) in Language Learning: Case of Tamazight

ABSTRACT: *No field today can do without the assistance of new technologies. Information technology invades all sciences to the great benefit of researchers. While these tools tend to facilitate analyses, they are not without flaws. The use of artificial intelligence in linguistic studies is very recent and struggles to establish itself in methodological practices. Although partially mastered, its use is becoming increasingly frequent. Nowadays, almost all researchers occasionally or regularly resort to machine translation through Google Translate. This website provides translations of texts in several different languages. Our interest through this writing is to review the linguistic constraints and particularities of Tamazight that could affect the reliability of translations and thus assess the possibility of introducing Tamazight among the official languages of Google Translate. Our objective is to suggest preliminary alternatives by leveraging the shortcomings observed in translations into other languages. We aim to highlight translation situations that escape artificial intelligence and require intervention and reformulation of the source or target sentence. Furthermore, we will present the advantages of machine translation in the language learning process compared to other traditional learning systems.*

KEYWORDS: Assisted/automatic translation, artificial intelligence, automatic language processing, Google translation, language learning

RÉSUMÉ : *Aucun domaine ne peut de nos jours se passer de l'assistance des nouvelles technologies. L'informatique envahit toutes les sciences au grand intérêt du chercheur. Si ces outils ont tendance à faciliter les analyses, elles ne sont pas sans failles. L'utilisation de l'intelligence artificielle dans les études linguistiques est très récente et peine à s'installer dans les pratiques méthodologiques. Bien que partiellement maîtrisée, son utilisation devient de plus en plus fréquente. De nos jours, presque tous les chercheurs ont occasionnellement ou régulièrement recours à la traduction automatique à travers Google traduction. Ce site web fournit des traductions de textes dans plusieurs langues différentes. Notre intérêt à travers cet écrit est de passer en revue les contraintes et particularités linguistiques du tamazight qui pourraient perturber la fiabilité des traductions et ainsi statuer sur l'éventualité de l'introduction de Tamazight parmi les langues officielles de Google translate. Notre objectif est de suggérer des alternatives en amant en se servant des insuffisances constatées à travers les traductions dans les autres langues. Nous visons à mettre en évidence les situations de traduction échappant à l'intelligence artificielle et nécessitant une intervention et une reformulation de la phrase source ou cible. Par ailleurs, nous exposerons les avantages de la traduction automatique dans le processus d'apprentissage des langues en l'opposant aux autres systèmes d'apprentissage classiques.*

MOTS-CLÉS : Traduction assistée/automatique, intelligence artificielle, traitement automatique des langues, Google traduction, apprentissage des langues

* Auteur correspondant : GUERCHOUH Lydia, guerchouh@yahoo.fr

ALTRALAG Journal / © 2024 The Authors. Published by the University of Oran 2 Mohamed Ben Ahmed, Algeria.

This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

Introduction

L'intervention de la machine dans les sciences sociales devient de plus en plus fréquente. Les études linguistiques sont assistées par différents logiciels et applications, destinées majoritairement au traitement automatique des langues. L'intelligence artificielle, quant à elle, peine à s'adapter à l'hétérogénéité et la dynamique linguistique qui ne compte qu'une timide introduction dans le domaine des sciences du langage. L'assistance de la machine est de plus en plus recherchée notamment en traduction visant l'accessibilité aux travaux scientifiques et la diffusion des résultats des recherches à travers le monde.

En effet, le brassage culturel et la tendance à l'effacement des frontières forcent l'évolution des moyens de communication exigeant rapidité et fiabilité dans la traduction impliquant des moyens automatiques qui réduirait l'intervention humaine coûteuse en temps et en argent. Plusieurs applications ont vu le jour dans le domaine de la traduction améliorant progressivement le rendu en fonction des particularités linguistiques et du volume du corpus disponible dans les différentes langues.

Notre intervention se penchera du côté de la traduction automatique via l'aspect de l'intelligence artificielle. Nous tenterons à travers une expérience dans la traduction des langues Français – Anglais – Arabe, de la possibilité d'intégration de Tamazight dans le moteur de recherche google traduction en répondant aux interrogations suivantes :

- A quels degrés les traductions automatiques sont-elles fiables ?
- Quelles sont les contraintes techniques et linguistiques dans le cas de l'introduction de Tamazight dans google traduction ?

Il convient, au préalable, de définir quelques concepts et mettre en évidence les oppositions nécessaires pour la compréhension des points qui suivront.

I. .1. Intelligence artificielle et traitement automatique des langues

Les deux concepts sont très souvent amalgamés et employés comme synonymes se distinguant de la linguistique informatique qui visent à comprendre les langues au moyen d'outils informatiques en opposition aux premiers dont le but est de mettre en place des outils informatiques pour assister l'humain dans le traitement des langues.

Le passage vers l'intelligence artificielle en linguistique était envisagé avec l'apparition du traitement automatique des langues (TAL). Ce dernier, qui était, autrefois, défini comme toutes applications servant à aider le linguiste dans ses analyses linguistiques s'est élargi pour rendre compte de l'intelligence artificielle ayant essentiellement débuté avec la traduction sur google. Toutefois, celle-ci se distingue progressivement du TAL dont l'objectif est l'assistance de l'humain et non son remplacement comme par exemple l'application servant à repérer les différentes catégories lexicales des unités employées dans un texte.

En effet, l'intelligence artificielle (I.A.) regroupe un ensemble de techniques mises en œuvre pour simuler et imiter l'intelligence et le mécanisme d'un cerveau humain. Ce sont, donc, des outils auxquels on confère des capacités d'analyse et de décision qui leur permettent de s'adapter intelligemment aux situations (à la manière d'un cerveau humain) en se servant de données déjà incluses. Les machines dotées d'I.A. sont à même d'adapter leurs résultats en fonction de différentes situations, d'actualiser ses informations allant même jusqu'à être capables d'interagir avec une personne humaine.

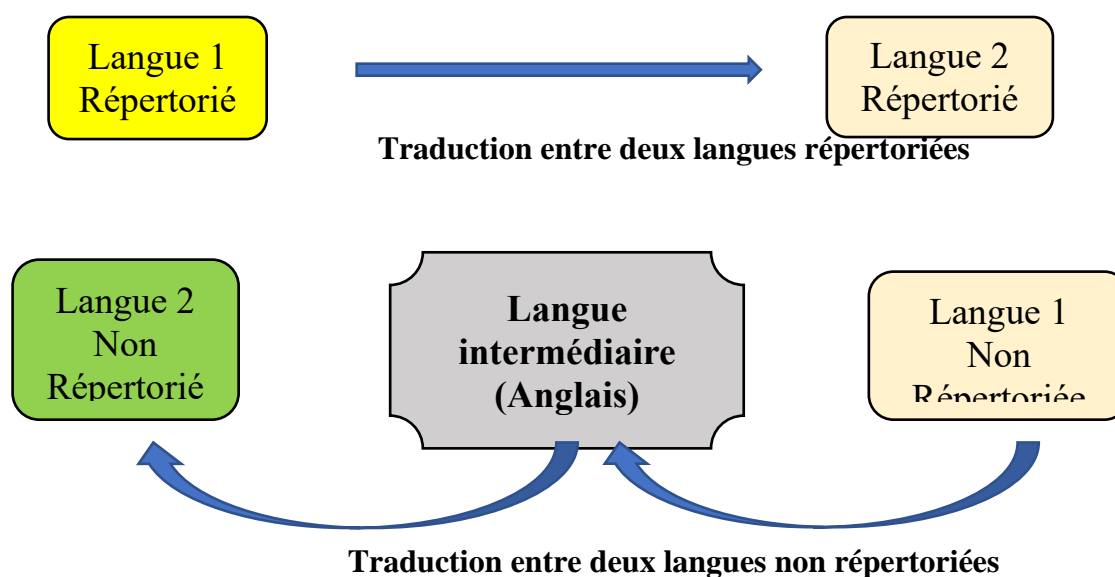
I.2. Traduction assistée et traduction automatique

Par traduction assistée, il est question d'utilisation de moyens informatique pour aider l'humain à traduire à son texte. C'est donc, lui qui fait la traduction en se servant, par exemple, d'une application terminologique lexicale qui va lui faciliter la tâche.

Quant à la traduction automatique, il s'agit là d'un outil informatique chargé d'établir la traduction à la place de l'humain dont l'intervention reste peu ou pas du tout requise. Ce concept rejoint, donc, celui d'intelligence artificielle.

I.3. Naissance et fonctionnement de Google Traduction

Les premiers essais de traduction automatique ont constitué le premier pas vers l'intelligence artificielle à travers la création de Google Traduction. Celui-ci, a commencé à traduire depuis et vers les six langues officielles de l'ONU : l'Anglais, le Français, l'Arabe, le Chinois, l'Espagnole et le Russe. Tous les documents et discours étaient alors instantanément traduits dans les six langues et constituaient une banque de données pour le moteur de recherche Google Traduction. Ce corpus qui s'enrichissait continuellement était la source des traductions automatiques proposées qui puise de ses structures syntaxiques et sémantiques.



La traduction a alors commencé avec ces langues dites répertoriées et pour traduire d'une langue non répertoriée à une autre langue non répertoriée, on se servait de l'anglais comme langue intermédiaire. Ainsi, la traduction est établie en référence aux structures composant le corpus à travers un algorithme qui proposera, pour toute traduction, la construction équivalente la plus fréquente parmi celles enregistrées :

Exemple : Lors de la traduction de la phrase suivante en anglais : « j'aime le lait »

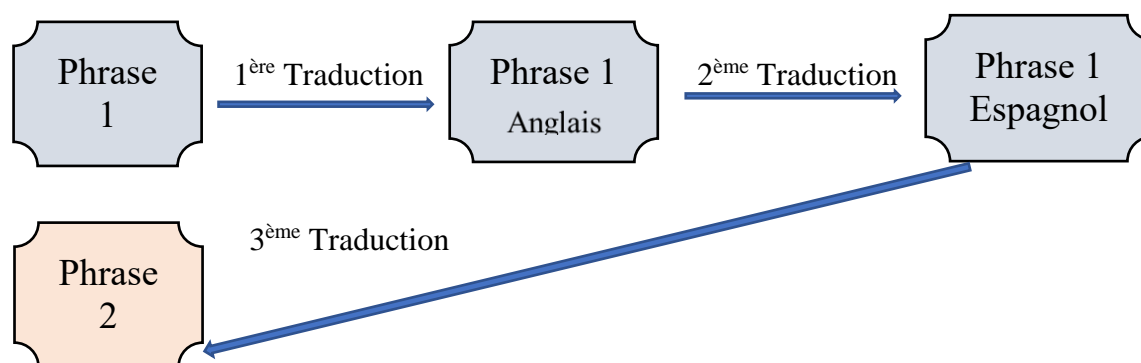
Le répertoire qui comporte plusieurs formes traduites sera réduit par l'algorithme à la forme la plus courante et la proposera sur l'interface :

« I love milk » *احب الحليب*
« I like milk » *احب شرب الحليب*
« I like drink milk »

Les progrès enregistrés dans Google Translate sont le résultat de deux facteurs principaux : d'une part, l'augmentation de la masse de données textuelles dites corpus disponibles sur internet et, d'autre part, la puissance de plus en plus importante de calcul des machines en filtrant des milliers de combinaison en un laps de temps pour en proposer la plus fréquente.

II. Les contraintes de la traduction automatique de et vers Tamazight

Il n'existe aucune traduction automatique fiable à 100%, ce pourquoi l'on parle plus de dégrossir et non traduire car l'intervention humaine demeure inévitable et souvent même recommandée. Aucun logiciel/application ne propose des traductions conformes même entre les langues dont les corpus sont les plus conséquents. Le cerveau humain reste le seul véritable traducteur lorsqu'il maîtrise les systèmes linguistiques concernés. Si la machine ne peut se référer qu'aux combinaisons répertoriées dans ses corpus, elle ne peut remplacer l'humain qui a la capacité d'adaptation, d'analyse et d'interprétation.



Si la première et la seconde traduction semblent respecter approximativement le sens de la phrase 1 en Français, le retour à la version française après deux traductions vers deux autres langues (l'Anglais et l'Espagnole), donne une phrase complètement différente de la phrase de départ qui semble être comprise et correctement traduite la première et la seconde fois.

Nous mettons, ici, l'accent sur les obstacles qui pourraient entraver la démarche de traduction en tamazight tout en comparant le degré de leur impact sur le texte traduit dans les autres langues.

II.1. L'ambiguïté sémantique

Outre les difficultés d'interprétation liées aux expressions figées, les constructions phrastiques rendent épineuses toute tentative de traduction.

II.1.1. L'homonymie et la polysémie

Il n'est nul doute que ces caractéristiques linguistiques constituent une barrière rigide pour les traducteurs automatiques notamment lorsque aucun indicateur syntaxique et/ou sémantique n'est présent pour orienter le sens de la phrase ou que plusieurs rejoignent un seul et même référent.

Imi yesea imi, anef-as i netta ad yehder
Une bouche il a une bouche, laisse-le parler (Construction incorrecte)
Puisqu'il a une bouche, laisse-le parler

Yewwet-it yerna yettru
Il l'a frappé a rajouter il a pleuré (construction incorrecte)
Il l'a frappé et a osé pleurer

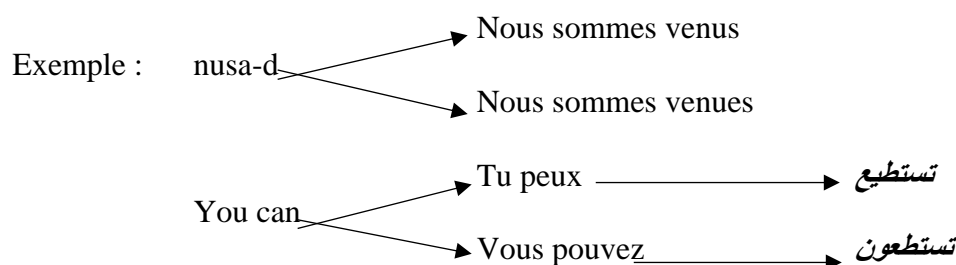
L'économie du langage étant à l'origine des phénomènes homonymiques et synonymiques, n'a épargné aucune langue et constitue, donc, une contrainte plus ou moins importante en fonction de la fréquence de ces cas. En Tamazight, l'homonymie ne semble pas être aussi contraignante que dans d'autres langues comme l'anglais où l'inventaire des homonymes est plus important. Ceci dit, la pauvreté du corpus, ne nous permet pas de lever l'ambiguïté de ces homonymies bien qu'elles soient en nombre limité : le verbe « yerna » serait traduit comme conjonction de coordination lorsqu'il est mis entre deux verbes (propositions) si le répertoire de cette langue était assez important pour que l'algorithme puisse en extraire les combinaisons probables.

Cependant, la polysémie très prépondérante en Tamazight constitue, parfois, un labyrinthe sémantique qu'il est, dans certains cas, impossible de résoudre.

Outre ces phénomènes que se partagent quasiment toutes les langues, la variation sémantique intra et inter dialectale peut perturber le sens de la traduction obtenue sans qu'aucun élément du co-texte n'intervienne pour lever l'ambiguïté.

II.1.2. L'ambiguïté de la référence personnelle

Bien que la majorité des langues soient construites autour de déclinaison permettant aisément d'identifier le référent personnel, il n'est pas le cas dans certaines langues. La traduction de et vers Tamazight qui, à l'instar de l'anglais, ne présente pas toujours des oppositions formelles, ouvre le champ de l'imprécision et donc de la traduction par fréquence et dont le résultat est réduit à une seule forme, généralement celle du masculin singulier la plus courante.



II.2. Contraintes grammaticales

Les contraintes d'ordre grammatical sont les plus compliquées à appréhender puisqu'elles ne peuvent être décodées que par le locuteur de la langue de manière manuelle.

II.2.1. La référence temporelle et modale

La langue amazighe étant une langue aspectuelle, la traduction des différents temps et modes verbaux devient un supplice qui détruit toute chronologie et relation dans une chaîne de verbes. Bien que

des compléments lexicaux peuvent constituer des alternatives à ses valeurs incluses dans le verbe, leur choix suppose une maîtrise profonde de la langue que l'informatique ne peut suggérer.

Exemple :

Il va venir	Ad d-yas
Il viendrait	ad d-yas

Aucune distinction ne peut être prise en charge dans la traduction en tamazight qui pourrait être, manuellement modalisée par une unité à part entière « ahat » (peut-être)

II.2.2. La référence au schème

Le schème est parfois l'indicateur principal du sens à traduire puisque, lorsqu'ils sont en distribution complémentaire, ils lèvent toute ambiguïté sémantique amenant à des traductions artificielles. En Tamazight, la répartition des schèmes est des plus irrégulière à plusieurs niveaux. La reconnaissance des aspects verbaux devient presque impossible pour les verbes à trois thèmes, notamment avec un aoriste nu (sans « ad ») qui est amalgamé à la forme du prétérit.

Exemple :

Win yebyan lyella, yekrez akal-is
Celui qui veut une bonne récolte, **a labouré** son champ (Construction incorrecte)
Celui qui veut une bonne récolte, **doit labourer** son champ

II.2.3. Le lexique et la terminologie

Contrairement aux langues ayant été intégrées dans Google traduction, Tamazight ne dispose pas encore de lexique référentiel couvrant tous les référents. Une contrainte capitale qui retardera, sans nul doute, son introduction dans la plateforme de traduction connectée.

Toutefois, à l'instar des autres langues y compris celles répertoriées, les jargons semblent constituer un obstacle dans la traduction automatique dans la mesure où, les répertoires cumulés relèvent en majorité du lexique ordinaire qu'il est quasiment improbable qu'une unité isolée soit proposée par l'algorithme à la place d'une unité fréquemment employée et dont le sens s'y rapproche.

Par ailleurs, si dans certaines langues, certaines parties des terminologies de spécialités peuvent être suggérées par google traduction, notamment celles relatives aux domaines médical et technologique qui sont les plus répertoriés, il est quasiment impossible d'obtenir un quelconque résultat avec la langue amazighe qui ne comporte, dans ces domaines, que des néologismes très récents et dont l'emploi est très rare et sporadique.

III. Les avantages de l'apprentissage des langues par le biais de l'intelligence artificielle

Pour nous permettre d'évaluer l'impact positif de la technologie sur le processus d'apprentissage des langues, nous avons opposé ce type d'apprentissage à l'apprentissage ordinaire (en salle de cours) afin de mettre en exergue les points forts de cette méthode d'acquisition des langues.

III.1. La disponibilité :

C'est le premier avantage accordé à la traduction numérique qu'elle soit via les applications lexicologique et/ou l'intelligence artificielle. En effet, l'accessibilité à ce moyen de traduction facilite et encourage même l'apprentissage des langues.

III.2. L'assistance dans la pratique orale :

Contrairement à l'apprentissage à travers les ouvrages ou l'enseignement classique, la traduction numérique ne fait pas uniquement dans l'équivalence graphique des unités mais permet d'apprendre la

prononciation correcte de ceux-ci et même la segmentation dans les situations d'assimilation des fragments de phrases, locutions, mots composés...

III.3. L'apprentissage adaptatif :

L'un des points faibles de l'apprentissage en classe notamment dans l'acquisition des langues étrangères, réside dans l'hétérogénéité des profils des apprenants. Les applications de traduction alimentées par l'intelligence artificielle permettent à chacun de suivre un rythme d'apprentissage individuel, de cibler les aspects de la langue dans lesquels il veut se perfectionner. En outre, certaines applications proposent même, sur la base de l'historique de recherche, une panoplie d'exercices adaptés au niveau et aux besoins spécifiques de l'utilisateur.

III.4. L'interactivité et l'autocorrection :

Grâce aux applications, l'apprenant peut s'exercer aux différents sens et significations des mots, à la synonymie ... en l'interrogeant à travers des traductions de segments adaptés à différents contextes. L'apprenant prend, donc, connaissance des contextes d'utilisation des unités mais aussi des variations lexicales. Par ailleurs, à plusieurs reprises, on nous a confié l'utilisation régulière de la traduction assistée non pour avoir l'équivalent lexical d'un mot mais uniquement pour s'assurer de son orthographe.

III. Conclusion :

Il est indéniable que google traduction constitue un outil de traduction non seulement efficace mais de plus en plus précis en fonction des langues. Après plusieurs essais avec diverses phrases, la traduction de et vers l'anglais semble la plus fiable y compris avec des tournures et des structures sémantiques non courantes. De sérieux problèmes de lexicologie sont à signaler en ce qui concerne la traduction de et vers la langue arabe dans la mesure où il y a très souvent un manque de précision dans l'emploi des mots qui ne tient que très rarement compte des nuances sémantiques et variations contextuelles. En dépit de la richesse lexicale de cette langue, la traduction ne semble pas cibler et cerner les distinctions sémantiques et google propose souvent des termes génériques qui détournent parfois le sens recherché et diminue d'autres fois, la pertinence de l'information à traduire.

Il est clairement démontré que la traduction de et vers la langue Tamazight ne peut être incluse dans google traduction dans l'immédiat et ce pour plusieurs raisons :

- En premier, le lexique répertorié dans cette langue est insuffisant pour rendre compte de tous les sens et significations à traduire : le corpus numérique constitué en ligne est très réduit et extrêmement pauvre.
- Deuxièmement, l'irrégularité de certaines règles grammaticales et morphologiques augmente les difficultés de l'algorithme à sélectionner les unités correctes à afficher et l'induit plus fréquemment à l'erreur.
- Troisièmement, l'absence de dictionnaires de référence réduit considérablement les chances d'introduction de Tamazight dans le système de traduction de google.
- Enfin, le point le plus prépondérant qui se dégage de cette petite analyse, nous renvoie au manque important des versions en tamazight des documents et ouvrages internationaux disponibles dans diverses langues incluses dans le moteur de recherche google. Cette étape est cruciale pour constituer le fond lexicologique et sémantique nécessaire au fonctionnement des algorithmes.

Références

- BALLARD Michel(1990),, *La Traduction Plurielle*, Presses Universitaires de Lille,France
- BOURQUIN Guy(1990), « Du transfert à la traduction : quelle(s) sémantique(s) ? », *T.A. Informations*, 31-1, Paris, ATALA, pp. 57-69
- BOURQUIN Guy(1985), « L'intérêt linguistique de la traduction automatique », in *La Recherche française par ordinateur en langue et littérature*, Charpentier, Colette, et David, Jean, eds, Genève-Paris, Slatkine-Champion, pp. 149-152.
- BOURQUIN Guy(1984),« Quel statut épistémologique donner à la traduction automatique », *Contrastes*, A4, Paris, ADEC, pp. 111-119.
- BOURQUIN-LAUNEY Marie-Claude(1984), « De la traduction automatique à la réflexion linguistique : la traduction de la préposition anglaise 'through' », *Contrastes*, A4, Paris, ADEC, pp. 81-99.
- DANLOS Laurence(1990), « Degré d'abstraction des représentations intermédiaires en traduction automatique, un exemple : Eurotra », *T.A. Informations*, 31-1, Paris, ATALA, pp. 25-37.
- DELAVENAY Emile(1968), « La traduction automatique », in *Le Langage*, encyclopédie de la Pléiade, Paris, NRF, pp. 758-771.
- François-Régis Chaumartin [Pirmin Lemberger](#) (2020) *Le traitement automatique des langues - Comprendre les textes grâce à l'intelligence artificielle*, Editions DUNOD
- FUCHS Catherine, et alii., *Linguistique et Traitements Automatiques des Langues*, Paris, Hachette – Supérieur.
- LOFFLER-LAURIAN Anne-Marie(1983), « Pour une typologie des erreurs dans la traduction automatique », *Multilingua*, 2-2, Amsterdam, Mouton, pp. 65-78.
- LOFFLER-LAURIAN Anne-Marie(1982), « Traduction automatique de textes techniques et analyse d'erreurs », *Contrastes*, A1, Paris, ADEC, 1982, pp. 45-58.
- LOFFLER-LAURIAN Anne-Marie(1983), « Traduction automatique et enseignement », *Revue de Phonétique Appliquée*, 66-68, Université de l'Etat à Mons, Belgique, 1983, pp. 86-102.
- OITET Christian et NEDOBEJKINE Nikolai(1985), « L'informatique au service de la linguistique – Illustration sur le développement d'un atelier de traduction automatisée », in *La Recherche française par ordinateur en langue et littérature*, Charpentier, Colette et David, Jean, eds, Genève-Paris, Slatkine-Champion, pp. 139-148.
- Pottier Bernard(1990). Linguistique et intelligence artificielle. In: *Langages*, 22^e année, n°87, 1987. Sémantique et intelligence artificielle, sous la direction de François Rastier. pp. 21-31.

Biographies des auteurs

Auteur 1 : Dr. Lydia GUERCHOUH, Maître de conférences A, Département de langue et culture amazighes, actuellement Doyenne de la Faculté des Lettres et des Langues après avoir occupé le poste de Vice Doyenne chargée des études et des questions liées aux étudiants. Également titulaire d'une Licence et d'un CAPA en Droit et d'une licence en biologie. Plusieurs publications nationales et internationales axées sur la syntaxe, la terminologie, la lexicographie/lexicologie ... et un ouvrage collectif édité par Cambridge Scholars Publishing en 2023. Cheffe de Projet PRFU, « Aménagement du lexique administratif et traduction des documents officiels en tamazight » jusqu'à Décembre 2026.

Auteur 2 : Prof. Dr. Mehmet Hakki Suçin est professeur de langue et littérature arabes à l'Université de Gazi, auteur et traducteur littéraire. Il a présidé la commission chargée d'élaborer des programmes d'études arabes basés sur le CECR en Turquie. Il a été l'un des membres du jury d'Arabic Booker en 2014. Il a reçu plusieurs prix de traduction en Turquie, tels que le prix de la meilleure traduction en 2016 pour sa traduction de *The Book of Love* de Nizar Qabbani et pour sa traduction de *Minor* d'Adania Shibli. Détails du roman en 2021. Dernièrement, reçu en 2022 Achievement Award pour tous ses travaux de traduction. Il a traduit vers le turc à la fois la littérature arabe classique comme al-Muallaqat, Al-Mutanabbi, Ibn Tufail, Ibn Hazm, Ibn al-Kalbi, et la littérature arabe moderne comme Kahlil Gibran, Adonis (4 ouvrages), Mahmoud Darwish (4 ouvrages).), Nizar Qabbani, Muhammad Bennis, Naguib Mahfouz, Ghassan Kanafani, Adania Shibli, Nouri Al-Jarrah, Najwan Darwish, Khulood Al Mualla et bien d'autres. Suçin a également traduit de la littérature turque vers des poètes soufis arabes tels que Yunus Emre, Ahmet Yesevi, parmi de nombreux autres poètes de l'ère moderne. Ses études se concentrent sur les études de traduction, la littérature arabe et l'enseignement de l'arabe aux locuteurs non natifs. Il travaille récemment à l'Université Gazi d'Ankara, en Turquie.