

الجمهورية الجزائرية الديمقراطية الشعبية  
République Algérienne Démocratique et Populaire  
وزارة التعليم العالي والبحث العلمي  
Ministère de l'Enseignement Supérieur et de la  
Recherche Scientifique

المدرسة العليا للإعلام الآلي بسيدي بلعباس ٨ ماي ١٩٤٥  
École Supérieure en Informatique  
-08 Mai 1945- Sidi Bel Abbès



icosnet

**THESIS**

To obtain the diploma of :**Master**  
Field : **Computer Science**  
Specialty :**Ingénierie des Systèmes Informatiques (ISI)**

**THEME**

---

**étude comparative entre les approches utilisées dans le sous-titrage**

---

présenté par :  
**Sadoun feriel**

Mr.Abdelhak Sifi (Icosnet)

*superviseur*

Mr.Khaldi belkacem

*superviseur*

Mr.chaib Souleymane

*Président*

Mr.Bekkouche Mohamed

*Examineur*

# Reconnaissance

Tout d'abord, je voudrais exprimer mes sincères remerciements à **Allah** pour m'avoir aidé et m'avoir donné. J'ai la persévérance et la motivation nécessaires pour faire ce travail. Mes remerciements et ma gratitude à mon patron **Dr. Sifi Abdelhak** Nous tenons à remercier le **Dr KHALDI Belkacem** de l'équipe de traitement du langage naturel et de contenu numérique pour ses précieux conseils et ses immenses connaissances tout au long du processus. Cette étude. Avoir un soutien thématique, persévérance, vision, intégrité et motivation J'étais profondément ému. Ce fut un grand privilège et un honneur de travailler sous la supervision thématique. Nous tenons également à remercier tous les juges qui ont accepté de passer leur temps à juger. mon boulot. Enfin, je tiens à remercier tous les professeurs et tous les employés de **ESI-SBA** .

# Résumé

Le sous-titrage est une technique liée aux contenus audiovisuels, notamment cinématographiques, consistant en l’affichage de texte au bas de l’image, lors de la diffusion d’un programme, comme un film. Cette technique qui a d’abord été utilisée pour le cinéma, a ensuite été transposée à la télévision, où elle peut concerner tous types de programmes comme les séries télévisées, les documentaires, les journaux télévisés, etc .

Elle s’applique désormais à tous les médias audiovisuels : DVD-Video, Internet, etc. Les sous-titres inter-langues sont des objets difficiles à appréhender théoriquement. Si l’on peut passer d’une langue à l’autre, c’est aussi passer de l’écoute à la lecture, du cinéma à l’écriture. Les caractères interlingues sont souvent cités pour considérer les sous-titres comme des traductions. Cependant, sa complexité interne suggère qu’il est plus qu’une simple traduction. En même temps, il est soumis à de fortes contraintes (exigences de brièveté, de lisibilité, de cohérence avec l’imagerie du film, etc.) qui le rendent moins précis que la traduction, moins long, et plus « équivalent » au texte original. La « sensibilité » sera plus faible. Cet ouvrage propose un examen théorique de la relation entre sous-titrage et traduction. À cette fin, nous utilisons la terminologie scientifique de base de la traduction pour le sous-titrage. Premièrement, le passage de l’oral au sous-titrage est étroitement lié à la distinction linguistique entre traduction et interprétation. Afin de pouvoir résumer les dialogues et résumer leur contenu, les sous-titres doivent les interpréter. Cela nous permet de souligner la relation étroite entre traducteur, interprète et interprète. Deuxièmement, les restrictions strictes imposées au sous-titrage nous amènent généralement à remettre en question la notion d’équivalence au cœur de la traduction.

**Mots-clés** : sous-titrage, transcription, traduction automatique, Natural Language Processing, Machine learning, Deep learning.

## summary

Automatic transcription and machine translation are two key technologies in the field of linguistics and communication.

Automatic transcription is the process by which programs or artificial intelligence systems automatically convert speech into written text. This technology is widely used in areas such as transcribing audio or video files, creating subtitles for media, medical transcription, and many other applications. It relies on speech recognition algorithms that analyze audio signals and convert them into written words.

Machine translation, on the other hand, is the ability of computer systems to automatically translate text from a source language to a target language without human intervention. Machine translators use natural language processing (NLP) algorithms to analyze the grammatical structure and meaning of sentences in the source language and then generate a translation in the target language. Examples of popular machine translation tools include Google Translate and DeepL, which are widely used.

These two technologies have advanced significantly in recent years thanks to advances in artificial intelligence, especially with the use of deep neural networks. These advances have improved the accuracy and quality of automatic transcriptions and translations, making them increasingly useful in a wide range of areas, including international communication, content localization, linguistic research, and information management.

In summary, automatic transcription converts speech into written text, while machine translation allows for the automatic translation of text from a source language to a target language, using artificial intelligence technologies to enhance the accuracy and speed of multilingual communication processes.

## ملخص

النص التلقائي والترجمة التلقائية هما تقنيتان رئيسيتان في مجال علم اللغة والاتصال. النص التلقائي هو العملية التي يتم فيها تحويل الكلام تلقائياً إلى نص مكتوب باستخدام برمجيات أو أنظمة ذكاء اصطناعي. تُستخدم هذه التكنولوجيا على نطاق واسع في مجالات مثل نقل الملفات الصوتية أو الفيديو، وإنشاء ترجمات لوسائل الإعلام، والنصوص الطبية، وغير ذلك. يعتمد ذلك على خوارزميات التعرف على الصوت التي تحلل الإشارات الصوتية وتحولها إلى كلمات مكتوبة.

أما الترجمة التلقائية، فهي قدرة الأنظمة الحاسوبية على ترجمة نص تلقائياً من لغة مصدر إلى لغة هدف دون تدخل بشري. يستخدم مترجمون آليون خوارزميات معالجة اللغة الطبيعية لتحليل البنية اللغوية والمعنى في الجمل في اللغة المصدر، ثم توليد ترجمة في اللغة المستهدفة. أمثلة على مترجمين آليين شهيرين مثل جوجل ترانسليت وDeepL.

شهدت هاتان التقنيتين تطوراً كبيراً في السنوات الأخيرة بفضل تقدمات الذكاء الاصطناعي، ولا سيما باستخدام الشبكات العصبية العميقة. هذه التقدّمات ساهمت في تحسين دقة وجودة النصوص التلقائية والترجمات، مما جعلها أكثر فائدة في مجموعة واسعة من المجالات، بما في ذلك الاتصال الدولي، وتخصيص المحتوى، وأبحاث اللغة، وإدارة المعلومات. باختصار، النص التلقائي يحول الكلام إلى نص مكتوب، بينما الترجمة التلقائية تسمح بترجمة نص تلقائياً من لغة مصدر إلى لغة هدف باستخدام تقنيات الذكاء الاصطناعي لتحسين دقة وسرعة عمليات الاتصال متعددة اللغات.