



Universidad  
de Alcalá

# **El traductor iFLYTEK: análisis y aplicación en el ámbito sanitario (español-chino)**

科大讯飞翻译机在医疗领域的应用及其特点分析（西汉）

**Máster Universitario en Comunicación Intercultural,  
Interpretación y Traducción en los Servicios Públicos**

**Presentado por:**

**D<sup>a</sup>. Zihan Sun**

**Dirigido por:**

**D<sup>a</sup>. María del Mar Sánchez Ramos**

**Alcalá de Henares, a 1 de junio de 2023**

## **AGRADECIMIENTO**

Antes de comenzar el estudio, me gustaría expresar mis más sinceros agradecimientos a todos aquellos que me han brindado ayuda y apoyo. Sin su orientación y compañía, habría sido muy difícil para mí terminar la redacción de esta tesis.

En primer lugar, quería agradecer con todo mi corazón a mis familiares, quienes siempre están a mi lado, brindándome soporte espiritual y económico. No importa en cualquier momento, estoy gratuíta por la motivación y el infinito amor de ellos.

Asimismo, me gustaría agradecer mi tutora, María del Mar Sánchez Ramos, quien me ha orientado y corregido en la redacción. Su paciencia, amabilidad y consejos han sido fundamentales para mi crecimiento.

Por último, es esencial agradecer a todas aquellas obras citadas y herramientas utilizadas en mi tesis, las cuáles han sido fuente de inspiración y me han ayudado a consolidar la base teórica para realizar este estudio.

Atentamente

## Índice

<b>RESUMEN</b> .....	5
<b>摘要</b> .....	6
<b>Introducción</b> .....	7
Estructura del estudio .....	7
Objetivo del estudio .....	8
<b>1. Traducción en el ámbito sanitario</b> .....	9
<b>1.1 Tipología de traducción</b> .....	10
1.1.1 Traducción no especializada.....	10
1.1.2 Traducción especializada.....	11
<b>1.2 Características del lenguaje médico</b> .....	12
1.2.1 Características del lenguaje médico español .....	12
1.2.1.1 Características sintácticas .....	12
1.2.1.2 Características morfológicas.....	14
1.2.2 Características del lenguaje médico chino.....	16
1.2.2.1 Características sintácticas .....	17
1.2.2.2 Características morfológicas.....	18
<b>1.3 Características y dificultades de traducción sanitaria en España y China</b> .....	20
<b>1.3.1 Características de traducción sanitaria</b> .....	20
<b>1.3.2 Dificultades de traducción sanitaria</b> .....	25
1.3.2.1 Falta de formación sistemática .....	25
1.3.2.2 Falta de código deontológico.....	26
<b>2. La traducción automática en contextos especializados</b> .....	28
<b>2.1 Traducción automática: definición y tipología de sistemas</b> .....	28
2.1.1 Definición de TA .....	28
2.1.2 Desarrollo de la TA.....	30
2.1.3 Tipología de sistemas .....	31
<b>2.2 La traducción automática en el ámbito sanitario</b> .....	41
<b>2.3 Evaluación de la TA</b> .....	45
2.3.1 Evaluación automática de la calidad de TA.....	47
2.3.2 Evaluación humana de la calidad de la TA.....	51
<b>2.4 Particularidad de traductor iFLYTEK</b> .....	59
<b>3. Muestra de datos y metodología</b> .....	61
<b>3.1 Descripción del estudio</b> .....	61
<b>3.2 Muestra de datos</b> .....	62
3.2.1 Creación de corpus .....	62
3.2.2 Traducción de corpus (español-chino).....	69
<b>3.3 Evaluación y Análisis</b> .....	70
3.3.1 Evaluación automática con BLEU .....	70
3.3.2 Evaluación humana .....	90

<b>4. Análisis y resultados</b> .....	<b>9 5</b>
<b>4.1 Características de la TA de iFLYTEK</b> .....	<b>9 5</b>
4.1.1 Traducción literal.....	9 5
4.1.2 Lenguaje Coloquial .....	9 6
4.1.3 Falta de correspondencia .....	9 7
4.1.4 Problema de forma .....	9 7
4.1.5 Falta de contexto .....	9 8
<b>4.2 Comentarios de métricas de evaluación de BLEU</b> .....	<b>9 8</b>
4.2.1 No se puede distinguir los errores. ....	9 9
4.2.2 Los parámetros no son fijos.....	9 9
4.2.3 Se limita al texto de referencia. ....	1 0 0
<b>Conclusión</b> .....	<b>1 0 0</b>
<b>Bibliografía</b> .....	<b>1 0 1</b>
<b>Anexos</b> .....	<b>1 0 7</b>
Anexo 1: Texto original en chino (de referencia).....	1 0 7
Anexo 2: Textos originales en español .....	1 0 9
Anexo 3: Textos traducidos al chino por iFLYTEK .....	1 1 2

## RESUMEN

En consonancia con el Instituto Nacional de Estadística (INE) de España, desde 1988, el número de inmigrantes chinos en España ha aumentado significativamente, pasando de 12.306 a 223.591 hasta el 1 de enero de 2022.<sup>1</sup> Además, la cantidad de turistas chinos que visitan a España ha incrementado, con 699.108 llegadas en 2019, siendo los principales destinos Cataluña, Madrid y Andalucía.<sup>2</sup> Aunque la pandemia de coronavirus había disminuido de cierta manera el número de inmigrantes y turistas, con la apertura de las políticas y el alivio de las restricciones de viaje, la cantidad sigue creciendo rápidamente. Es por esto que ha generado una gran mayor necesidad de traducción en el ámbito sanitario.

No obstante, la traducción humana tradicional no puede satisfacer la demanda de los usuarios. Como líder de empresas de tecnología de traducción, iFLYTEK ha desarrollado una serie de traductores que cuenta con funciones de traducción e interpretación en tiempo real, reconocimiento de dialectos y acentos, traducción e interpretación sin conexión de internet, terminologías en campos específicos, conocimientos profesionales, etc. El objetivo de esta tesis consiste en analizar la aplicación del traductor iFLYTEK en el ámbito sanitario y explorar sus ventajas y limitaciones.

Para realizarlo, se ha creado un corpus bilingüe concentrándose en la tuberculosis como ejemplo y se traduce el texto español al chino utilizando iFLYTEK. Además, se aplican dos métricas: BLUE (Bilingual Evaluation underestudy) y TAUS (Translation Automation User Association) para realizar la evaluación automática y la evaluación humana de este texto traducido. Esta metodología requiere una base teórica:

- 1) La definición de traducción especializada
- 2) Características del lenguaje médico español y chino
- 3) Rasgos y dificultades de traducción sanitaria en España y China
- 4) La tipología de traducción automática
- 5) Evaluación automática y evaluación humana

En resumen, el análisis de la calidad de traducción muestra que el iFLYTEK tiene varios problemas en la traducción en el ámbito sanitario. Dada la capacidad limitada de la evaluación humana, este estudio toma como ejemplo el tema de la tuberculosis. Por lo tanto, aún queda mucho camino por recorrer para la investigación de la aplicación del traductor iFLYTEK en el ámbito sanitario.

---

<sup>1</sup> [https://www.ine.es/prensa/pad\\_2022\\_p.pdf](https://www.ine.es/prensa/pad_2022_p.pdf)

<sup>2</sup> <https://www.tourspain.es/eu-es/ConocimientoTuristico/FichasEjecutivas/Ficha%20Ejecutiva%20China%20Mayo%202020.pdf>

**Palabras clave:** traducción automática, iFLYTEK, evaluación humana, evaluación automática, ámbito sanitario

## 摘要

根据西班牙国家统计局（NIE）数据显示，自 1998 年以来，截至 2022 年 1 月 1 日，居住在西班牙的中国移民数量从 12036 人增加到 223591 人。2019 年，中国游客入境西班牙次数高达 699108 次，其中主要目的地是加泰罗尼亚，马德里和安达卢西亚。尽管受疫情影响，这两年游客人数有所下降，但随着各国政策的放宽和疫情的缓解，移民西班牙的中国人以及前往西班牙旅游的人数仍在迅速增长。因此，在医疗领域存在着很大的翻译需求。

然而，传统的人工翻译无法完全满足这种需求。作为翻译的龙头企业，科大讯飞开发的翻译机具有实时翻译、方言翻译、口音识别、离线翻译，配有专业领域术语库及专业知识等多种功能。本论文旨在探讨科大讯飞翻译机在西班牙语和中文之间医疗领域的应用，以及其特点优劣。

为此，本文以结核病为例，创建了双语语料库，然后应用科大讯飞对文本进行西班牙语到中文的翻译。并使用 BLEU(双语评价研究)和 TAUS（翻译自动化用户协会）分别对译文进行自动评价和人工评价，对比两种评价方式的特点，分析翻译文本的质量。应用此研究方法需要足够的理论支撑：

- 1) 定义特定领域的翻译
- 2) 分析西班牙语和中文医学语言的特点
- 3) 总结两个医疗领域翻译的特点以及遇到的难点
- 4) 分析机器翻译系统的分类
- 5) 研究两种评价译文的方式
- 6) 介绍科大讯飞翻译机的特点

最后，通过文本翻译质量的分析，科大讯飞在医疗领域的翻译仍然存在许多问题。鉴于人工评价的能力有限，本文对于医疗文本翻译的研究以结核病为例，研究范围不够广泛，多样。因此，对于科大讯飞的翻译功能仍有待进一步研究。

**关键词：**机器翻译，科大讯飞，人工评价，机器评价，医疗领域

## **Introducción**

Debido a la demanda de la realidad, es necesario explotar y evaluar el potencial de este traductor en la traducción sanitaria. Este traductor es conocido en China, es que esta máquina iFLYTEK no es un sitio web ni una aplicación, sino una máquina de mano que se puede llevar consigo. Según la introducción de la empresa, este traductor es capaz de satisfacer la necesidad diaria, así como la necesidad en los campos específicos como campo jurídico, médico, comercial, administrativo, etc. Entonces, en consonancia con la importancia de la sanidad para la población, esta tesis toma objetivo analizar las características y evaluar la calidad de la traducción en el campo sanitario.

### **Estructura del estudio**

Esta tesis se divide en dos partes y cinco capítulos. La estructura es muy clara y fácil y está compuesta por el marco teórico y la metodología.

En cuanto al marco teórico, se explican las definiciones de términos relacionados en esta tesis, se exponen las características del lenguaje médico, y se describe la traducción automática.

1) En el primer capítulo, se introduce la tipología de traducción, que se clasifica en traducción no especializada y traducción especializada. El motivo de clasificarla consiste en exponer que la traducción en el ámbito sanitario como la traducción especializada requiere conocimientos específicos y terminologías profesionales. Además, se resumen las características del lenguaje médico español y chino. En el lenguaje médico español, se utilizan mucha voz pasiva, abuso del gerundio, solecismos, terminología, extranjerismos innecesarios y fijos amigos, mientras que, en el lenguaje médico chino, se aplican con frecuencia la voz pasiva, expresión fija, expresión de normalización y de prudencia. A lo siguiente, se interpretan las características y dificultades de traducción sanitaria en España y en China. Los rasgos consisten en amplitud de campo, situaciones comunicativas y géneros textuales, así como terminologías y culturas. Las dos principales dificultades son falta de formación sistemática y falta de código deontológico tanto en España como en China.

2) En el segundo capítulo, se describe la traducción automática en contextos especializados. Antes que nada, se explican la definición de la traducción automática y el desarrollo de la historia de TA. Los sistemas de traducción automática se dividen en cinco tipos: sistemas de traducción automática basados en reglas, sistemas de traducción automática basados en estadística, sistemas de traducción automática basados en ejemplos, sistemas de traducción automática híbridos y sistemas de traducción automática basados en redes neuronales. Después, se proponen los estudios actuales dedicados a la traducción automática en el ámbito sanitario. Existen dos maneras para hacer la evaluación de traducción: evaluación de la TA, y

evaluación humana. Se explican de manera detallada los modelos de las métricas.

En cuanto a la metodología, se crea un corpus para hacer el análisis y evaluación. En este proceso, aplican la herramienta de AntConc, ROST, y Notepad ++ para establecer un corpus bilingüe. Después, se traduce el corpus español con el traductor iFLYTEK. Entonces, se evalúa con las métricas de BLEU. De esta manera, se compara la similitud de texto traducido con el texto chino de referencia. En base a lo anterior, se analizan el resultado y cada zona marcada en color rojo para definir la calidad de la traducción y la de la evaluación automática. A continuación, se evalúa la traducción por mí misma. Se encuentran los errores y se clasifican por diferentes categorías: errores graves, errores importantes, errores poco importantes, errores neutros y menciones excelentes. Para realizarlo, se aplica el modelo de TAUS (Translation Automation User Society) a fin de realizar esta evaluación. Por fin, en base al resultado de evaluación y la calidad de traducción, se resumen las características de la TA de iFLYTEK y se dan comentarios de métricas de evaluación de BLEU.

### **Objetivo del estudio**

Este estudio se concentra en el ámbito sanitario y se espera que, por medio de esta tesis, se pueda alcanzar los siguientes objetivos:

- 1) Aclarar las características del lenguaje médico español y chino para facilitar la traducción en el ámbito médico. Las estructuras de la expresión médica y los términos en lenguaje médico de los dos idiomas son diferentes, por lo que con estas características los traductores pueden tener un concepto de los rasgos y los puntos a que deben prestar atención antes de traducir y pueden prepararlos de antemano.
- 2) Identificar las características y dificultades de traducción sanitaria en España y China, que está encaminado a señalar la dirección a que los traductores pueden dirigirse. Si quiere mejorar la situación laboral y profesional de la traducción sanitaria, hace falta establecer formación y código étnico por medio de la colaboración de diferentes categorías.
- 3) Especificar las dos maneras de evaluar la traducción automática. Se introducen diferentes modelos y métricas actuales de evaluación automática y evaluación humana. En base a estos, se proporcionan referencia para aplicarlos en la realidad.
- 4) Comparar diferentes sistemas de traducción automática. Las TA basados en diferentes sistemas cuentan con sus propias ventajas y limitaciones. Se pueden elegir la manera más adecuada de acuerdo con la necesidad y la tipología del corpus.
- 5) Analizar las características de la traducción automática realizada por iFLYTEK a fin de saber las limitaciones que pueden mejorar y los beneficios que pueden aprovechar.
- 6) Explicar las herramientas utilizadas en el proceso de establecer el corpus. Es fácil relativamente establecer un corpus monolingüe, pero es necesario segmentar palabras chinas y transformar la codificación en caso de corpus bilingüe.
- 7) Comentar las dos maneras de evaluación automática y evaluación humana para que puedan elegir una adecuada según el motivo y la longitud de corpus.
- 8) Proporcionar corpus relacionados con la tuberculosis para que los chinos pueden consultar sobre la información relativas sobre los síntomas, los tratamientos, las pruebas, las vacunas, la prevención, etc.



## 1. Traducción en el ámbito sanitario

Antes que todo, la sanidad es un concepto muy amplio en que se incluyen los dos objetos imprescindibles, los proveedores de la atención sanitaria y los receptores de la misma (la población a la que atiendan), además dispone de una naturaleza de servicios públicos. Como se exponen las definiciones de la sanidad en diferentes diccionarios:

La sanidad es un conjunto de servicios organizados por una comunidad, de carácter preventivo, terapéutico o de rehabilitación, con la finalidad de alcanzar un nivel óptimo de salud tanto privada como pública en el conjunto de la comunidad correspondiente. Sin.: salud pública. (Diccionario: Real Academia Nacional de Medicina de España, 2023)<sup>3</sup>

*医疗卫生也称为医疗卫生事业或医疗卫生服务，包括国家与社会为保障和提高人民的健康水平、诊治疾病而建立的法制体系、组织体系、服务体系和服务过程等。医疗卫生是公共卫生和医疗服务的统称，涉及社会公共卫生服务、医疗服务、健康促进服务以及与这些服务相关的保障体系、组织管理和监督体系等。  
(百度百科, 2023)<sup>4</sup>*

Traducción propia: La sanidad es un conjunto de empresas médicas y sanitarias o servicios médicos o de salud, incluidos el sistema legal, el sistema organístico, el sistema de servicios establecido por el estado y la sociedad para garantizar y mejorar la salud de la población, diagnosticar y tratar enfermedades. Es un término general para la salud pública y los servicios médicos, que incluye los servicios sociales de salud pública, los servicios médicos, los servicios de promoción de la salud y los sistemas de seguridad, gestión organizativa y sistemas de supervisión relacionados con estos servicios. (Enciclopedia Baidu, 2023)

De acuerdo con John. J. Hanlon (2014), de nivel general, la salud pública se puede resumir en dos campos, uno es el campo: los problemas ambientales, y otro, el mercado de los servicios de salud.<sup>5</sup>

En cuanto a la traducción, es un proceso para transmitir un mensaje de un idioma a otro idioma, por esto podemos saber que los traductores desempeñan su papel para eliminar la barrera y dificultad de comunicación entre dos partes provenientes de diferentes culturales y que hablan diferentes idiomas. Como se exponen las opiniones de diferentes autores:

La traducción es un proceso interpretativo y comunicativo consistente en la reformulación de un texto con los medios de otra lengua, que se desarrolla en un contexto social y con una finalidad determinada. (Amparo Hurtado Albir, 2001)<sup>6</sup>

---

<sup>3</sup> Diccionario: Real Academia Nacional de Medicina de España, última citada: 2023

<sup>4</sup> Enciclopedia Baidu, última citada: 2023

<sup>5</sup> Hanlon, J. J., (2014). La filosofía de la salud pública. Revista Cubana de Salud Pública, 40(1), . ISSN: 0864-3466

<sup>6</sup> Albir, A. H. (2001). Traducción y traductología / Translation and translation studies: Introducción a la traductología /

翻译是把一种语言表达的信息用另外一种语言表达出来。(蔡毅, 2003)<sup>7</sup>

Traducción propia: La traducción está destinada a expresar información expresada en un idioma a otro idioma. (Yi Cai, 2003)

## 1.1 Tipología de traducción

Existen diferentes maneras para clasificar la tipología de traducción. De acuerdo con Amparo Hurtado Albir (2001), la traducción se divide en catorce diferentes tipos según las modalidades de traducción: traducción escrita, traducción a la vista, interpretación simultánea, interpretación consecutiva, interpretación de enlace, susurrado, doblaje, voces superpuestas, subtitulación, traducción de programas informáticas, traducción de productos informáticos multimedia, traducción de canciones, supra titulación musical y traducción icónico-gráfica.

Cabe destacar la importancia de saber la tipología de textos para los traductores, es que diferentes tipos de documentos llevan consigo diferentes rasgos y niveles de dificultades. En el proceso de preparación, los traductores tienen acceso a buscar textos paralelos o textos de pertinentes temas a fin de facilitar y mejorar la calidad de su traducción. Por lo consiguiente, según la opinión de Albir (2001), de acuerdo con la tipología de los textos, la traducción está clasificada por traducción de géneros especializados y traducción de géneros no especializados.

### Tabla 1

*Tipos de traducción (Amparo Hurtado Albir, 2001, p.59)*

<b>Ámbito marcados por el campo</b> → Traducción de géneros especializados Tipo de campo: técnico, científico, jurídico, económico, religioso, etc.→Traducción técnica, científica, jurídica, etc.
<b>Ámbito no marcados por el campo</b> → Traducción de géneros no especializados Traducción de géneros literarios, publicitarios, periodísticos, etc.→Traducción literaria, traducción publicitaria, traducción periodística, etc.

### 1.1.1 Traducción no especializada

Según la tabla 1, la traducción no especializada no se limita a un campo específico sino de naturaleza general. Los textos traducidos pueden tener diferentes modos como modo “oral, escrito, audiovisual, icónico-gráfico” (Albir, 2001) y también se pueden traducir por medio de variadas modalidades de traducción, como “traducción escrita, traducción a la vista, doblaje, interpretación, simultánea, etc.”. (Albir, 2001, p.62)<sup>8</sup>

---

Introduction to Translation. Catedra Ediciones.

<sup>7</sup> 蔡毅, (2003), 关于翻译理论中的几个基本概念问题, 中国俄语教学, 第22卷第1期 (Yi Cai, (2003), Sobre varias cuestiones conceptuales básicas en la teoría de la traducción, Enseñanza del ruso en China, Volumen 22, Número 1)

<sup>8</sup> Albir, A. H. (2001). Traducción y traductología / Translation and translation studies: Introducción a la traductología / Introduction to Translation. Catedra Ediciones. P62.

### 1.1.2 Traducción especializada

La traducción especializada generalmente está dedicada a un ámbito específico, tales como: médico, técnico, educativo, administrativo, económico, comercial, político, histórico, etc. No obstante, desde una perspectiva de concepto amplio, de acuerdo con Silke Buthmann (2013), la “vinculación de la traducción especializada con los textos especializados resulta bastante restrictiva”, es que en otros ámbitos que no se aplican “lenguaje de especialidad ni textos especializados” también se utiliza este tipo de traducción. En consecuencia, la comprensión de las características y la aplicación de traducción especializada es significativo para los traductores.

Según María Teresa Cabré (1993), los lenguajes de especialidad son “subconjunto del lenguaje general caracterizados pragmáticamente por tres variables: la temática, los usuarios y las situaciones de comunicación.” Es por ello que podemos enterarnos de que el lenguaje de un campo especializado tiene su propia tema, interlocutores y situaciones de aplicación, lo que conduce a diferentes rasgos, diferentes terminologías, diferentes conocimientos, diferentes competencias, diferentes niveles de dificultad para la traducción de los textos.

**Tabla 2**

*Características de la traducción de textos especializados (María Teresa Cabré 1993)*

CARACTERÍSTICAS DE FUNCIONAMIENTO TEXTUAL	COMPETENCIAS REQUERIDAS DEL TRADUCTOR	
Importancia del campo temático	Conocimientos temáticos	CAPACIDAD PARA DOCUMENTACIÓN
Terminología específica	Conocimientos de terminología	
Géneros característicos	Conocimiento de los géneros característicos	

Entonces, el tema de esta tesis se concentra en el ámbito sanitario, sin duda, es un campo muy específico. Así, para un traductor médico o sanitario, es imprescindible disponer de conocimientos temáticos, conocimientos de terminología y conocimiento de los géneros característicos. Además de estos, igualmente cuenta mucho la capacidad para documentación.

Explicado de manera concreta, según Bianca Vitalaru (2022)<sup>9</sup>, para un traductor sanitario hace falta disponer de los conocimientos temáticos, lingüísticos, terminológicos, culturales y recursos, así como las habilidades como comprender lo mensajes, saber utilizar diferentes registros, saber redactar, saber documentarse para aplicar ciertas técnicas para encontrar

<sup>9</sup> Bianca Vitalaru (2022), PPT, traducción y documentación. rev.pdf. p.10, Traducción especializada en el ámbito sanitario, Universidad de Alcalá

soluciones para los problemas y dificultades de traducción, saber utilizar los recursos disponibles para solucionar los problemas de traducción y saber traducir para aplicar ciertas técnicas para transmitir el mensaje adecuadamente.

Asimismo, en cuanto a los conocimientos de géneros característicos, de acuerdo con Bianca Vitalaru (2022), en base a los emisores, receptores, niveles de dificultades de los documentos textuales, los textos sanitarios se pueden dividir en los siguientes tres tipos: a). Documentos informativos y documentos de libre acceso para pacientes. b). Instrucciones o recomendaciones y documentos que el personal sanitario hace llegar al paciente. c). Documentos utilizados por el personal sanitario.

Como consecuencia, aprender distinguir la tipología de textos sanitarios nos ayuda tener un concepto de la dificultad y los requisitos de la traducción, puesto que cada tipo cuenta con diferentes características e intenciones y la clave de distinguirlo es localizar su emisor y receptor. En cuanto a los dos primeros tipos, desde hospital o el personal sanitario para pacientes, el uso de palabras es más sencillo y básico, mientras la estructura es menos complicada con terminología menos profesional y formal. Al contrario, con respecto al tercer tipo, documentos solo entre el personal sanitario, podemos imaginar que se trata de un texto con muy alta profesionalidad, especialidad y formalidad, es que solo están dirigidas a los expertos y todos ellos tienen conocimiento médico de alto nivel.

## **1.2 Características del lenguaje médico**

Como explicaron Rafael Aleixandre-Benavent, Francisco Jesús Bueno Cañigral y Loudres Catelló Cogollos (2017), “el lenguaje sanitario, al igual que todos los lenguajes científicos y técnicos, debe construirse sobre 4 principios básicos: la precisión, la corrección, la claridad y la concisión.” Es por eso que el lenguaje sanitario requiere conocimientos sobre la estructura, la sintaxis, las terminologías, el estilo, la estructura, etc. El lenguaje sanitario es una especialidad muy profesional, por lo que es imprescindible comprender las características a fin de la traducción sea comprensible para el destinatario.

### **1.2.1 Características del lenguaje médico español**

#### **1.2.1.1 Características sintácticas**

Antes que todo, se explican la definición y el significado de sintaxis según autores y diccionario:

La sintaxis consiste en parte de la gramática para estudiar el modo en que se combina las palabras y los grupos que estas forman para expresar significados, así como las relaciones que se establecen entre todas esas unidades. (Diccionario: Real Academia Española, 2023).<sup>10</sup>

---

<sup>10</sup> Diccionario: Real Academia Española, 2023

句法的研究对象是句子。句法是语法的一部分，它根据被研究语言的固定用法论述表语、修饰语和其他词的关系。句子的结构方式，表示其在句子中相互关系的词形式的排列。（百度百科）<sup>11</sup>

Traducción propia: La sintaxis consiste en un estudio de los componentes de una oración y su orden. El objeto de estudio de la sintaxis son las oraciones. La sintaxis es una parte de la gramática que se dedica a estudiar la relación de predicados, modificadores y otras palabras de acuerdo con el uso fijo del idioma. La forma en que se estructura una oración es la disposición de las ordenes de las palabras que indican su relación entre sí en una oración. (Enciclopedia Baidu, 2023)

### **Voz pasiva**

De acuerdo con Aleixandre Benavent y Amador Iscla (2002), debajo de la influencia de palabras inglesas, el lenguaje médico español se caracteriza por el abuso de voz pasiva, “con una construcción con el verbo auxiliar “ser”, o de voz pasiva refleja con la partícula “ser””. Desde el punto de gramática, no existen errores, pero en el campo de la ciencia médica, es mejor usar un objeto más claro y concreto. Es aconsejable eliminar el uso de voz pasiva innecesaria en la descripción de casos clínicos y en los resultados de pruebas.

### **Abuso del gerundio**

Además, de acuerdo con Aleixandre Benavent y Amador Iscla (2002), en los documentos sanitarios existe otro fenómeno muy común, que se usa con frecuencia el gerundio, aunque en ese contexto no es muy adecuado. Desde el punto de vista gramatical, el uso de gerundio es para indicar una acción simultánea con el verbo principal. Pero en muchas oraciones, se aplica gerundio en caso de no existir dos acciones simultáneas. Como exponen un ejemplo Benavent y Iscla (2002), en la situación siguiente: “en el cáncer de cérvix, el tiempo laparoscópico incluyó la preparación de los espacios vesicales, realizándose posteriormente la histerectomía”, el uso de “realizándose” no es adecuado, es que las dos acciones no suceden simultáneamente, sino una acción es más posterior.<sup>12</sup>

### **Solecismos**

En consonancia con Benavent y Iscla (2002), “el solecismo es un vicio de dicción consistente en emplear incorrectamente una expresión o en construir una frase con sintaxis incorrecta. Hay tres clases de solecismos o errores de sintaxis: de concordancia, de régimen y de construcción.” El ámbito sanitario es un campo relacionado directamente y estrechamente con la vida y salud de la población, por lo que el riguroso uso de palabras, gramática, construcción, estilo, terminología es un requisito indispensable.

---

<sup>11</sup> Enciclopedia Baidu, 2023

<sup>12</sup> Aleixandre Benavent, R. y Amador Iscla, A. (2002) “Problemas del lenguaje médico actual (III): Gramáticas y estilo” en Panace@, 11(1) Valencia: Tremédica.18-23

### 1.2.1.2 Características morfológicas

Antes que nada, se explica la definición de la morfología de acuerdo con diferentes autores y diccionario:

*语法分为句法和此法两个部分。句法研究的句子的内部结构，以词作为基本单位；词法研究的是词的内部结构，以语素为基本单位。（朱德熙《语法讲义》，1982. p25）*

Traducción propia: La gramática se divide en dos partes: sintaxis y morfología. La sintaxis estudia la estructura interna de las oraciones, con las palabras como unidad básica. La morfología es el estudio de la estructura interna de las palabras, con morfemas como unidad básica. (Zhu, Deixi, 1982. p25)<sup>13</sup>

Parte de la gramática que estudia la estructura de las palabras y de sus elementos constitutivos. (Diccionario: Real Academia Española, 2023) <sup>14</sup>

### Terminología

En primer lugar, como se menciona lo anterior, aprender un lenguaje específico requiere conocimientos profesionales. En el uso de terminologías médicas, aparecen con frecuencia errores de acentuarse. Son diferencias muy sutiles, pero con las se puede causar malentendidos e incluso no se puede entender el texto.

**Tabla 3**

*Términos que a veces se acentúa incorrecta (Aleixandre Benavent y Amador Iscla 2002)*

<b>Correcto con acento</b>	<b>Incorrecto</b>
Acené	Acene
Aldehído	Aldehido
Alérgeno	Alergeno
Apoplejía	Apoplejia
Diátesis	Diatesis
Glucólisis	Glucolisis
Hiperémesis	Hiperemesis
Hipocondría	Hipocondra
Maléo	Maleo
Nucléolo	Nucleolo
Ósmosis	Osmosis
Roséola	Roseola

<sup>13</sup> 朱德熙《语法讲义》，北京：商务印书馆，1982. p25 (Zhu. D. X., Conferencias de gramática, Commercial Imprenta de Beijing, 1982. p25)

<sup>14</sup> Diccionario: Real Academia Española, última citada: 2023

Táctil Zóster	Tactil Zoster
<b>Correcto sin acento</b> Electrodo Endocrino Epiglotis Estadio Estasis Homeostasis Intervalo Libido Miligramo Pediatria Periferia Píamadre Psiquiatre Radar Rubeola	<b>Incorrecto</b> Eléctrodo Endécrino Epíglotis Estadío Éstasis Homeóstasis Intérvulo Líbido Milígramo Pedíatra Perifería Píamadre Psiquíatre Rádar Rubéola

### Extranjerismos innecesarios

Debido a diferentes niveles de desarrollo de la ciencia médica, el fenómeno de extranjerismo se ve con frecuencia en el lenguaje médico. Según Rafael Aleixandre-Benavent, Francisco Jesús Bueno Cañigral y Lourdes Catelló Cogollos (2017), “para que un vocablo nuevo sea aceptado debería cumplir 2 condiciones: la primera, esencial, es que sea necesario, por no existir ninguna palabra equivalente en nuestra lengua; la segunda es que se adapte a las normas lingüísticas del español.”<sup>15</sup>

Deberíamos seguir las dos condiciones para evitar el extranjerismo innecesario en el lenguaje médico. Sin embargo, las palabras extranjeras aparecen más frecuentemente en nuestra vida. Aquí no solo se limitan a términos inglés, tales como abstract, fatal, feedback, host, lobar, ect. También existen muchos términos de otros idiomas.

#### Tabla 4

*Ejemplos de términos procedentes de otros idiomas (Aleixandre Benavent y Amador Iscla 2001)*

Lenguas de origen	Vocablos actuales
Alemán	bismuto, bloqueo, cinc, cuarzo, estribo, éster, mastocito, vaselina
Áeabe	álcali, alcohol, alquimia, azúcar, elixir, jarabe
Francés	bocio, chanco, debutar, drenaje, ,masaje, pipeta, reservorio, tisular,

<sup>15</sup> Benavent. R. A. , Bueno Cañigral F. J., Cogollos L. C., Rafael Aleixandre-Benavent,\*, Francisco Jesús Bueno Cañigral y Lourdes Castelló Cogollos, (2017), Características del lenguaje médico actual en los artículos científicos, Educ Med. 18(Supl 2):23-29

Griego	viable acné, afta, apoplejía, carcinoma, dermatología, diarrea, eritema, herpes, lepra, psicología, púrpura, síncope, síndrome
Hebrero	abad, fariseo, sábado
Holandé	droga, escorbuto, esprue
Idiomas africanos	kwashiorkor, mosca, tse-tse, uabaína
Idiomas asiáticos	agar, beriberi, bezoar, kala-azar, tabú, tatuaje, té
Inglés	bypass, cóctel, comité, láser, escáner, shock
Italiano	batería, belladona, influenza, ,alaria, pelagra, petequia
Latín	absceso, célula, erupción, forúnculo, nevo, piel, tiña, uña, verrugas, virus
L. amerindias	cacao, cobaya, cumarina, guanina, ipecacuana
L. prerromanas	izquierda, pestaña, sabañón
Portugués	albinismo, fetichismo, sarpullido

## Falsos amigos

Según Benavent, Cañigral y Loudres Catelló Cogollos (2017), “los falsos amigos surgen por errores en la traducción desde otras lenguas, debido al parecido entre la palabra extranjera y alguna del español.”<sup>16</sup> Es un problema muy común y muy fácil de ignorarlo. Por un lado, las dos palabras son muy parecidas en la estructura léxica, lo que hace al traductor o usuario confundir los significados. Por otro lado, como se relata en el texto anterior, para las palabras españolas, existe extranjerismos innecesarios, entonces, es fácil reconocer los falsos amigos como parte de extranjerismos.

### Tabla 5

*Ejemplos de defectos de traducción por falsos amigos ( (Aleixandre Benavent y Amador Iscla 2001)*

Términos en inglés	Falso amigo	Traducción concreta
actual, actually	actual, actualmente	real, realmente
application	aplicación	solicitud
bull	bull	ampolla, vesícula
condition	condición	estado, afección, enfermedad
dramatic	dramático	impresionante, espectacular, considerable
pathology	patología	anatomía, patológica
protrusion	protrusión	protuberancia, saliente
severe	severo	grave, intenso, fuerte

## 1.2.2 Características del lenguaje médico chino

<sup>16</sup> Benavent. R. A. , Bueno Cañigral F. J., Cogollos L. C., Rafael Aleixandre-Benavent,\*, Francisco Jesús Bueno Cañigral y Lourdes Castelló Cogollos, (2017), Características del lenguaje médico actual en los artículos científicos, Educ Med. 18(Supl 2). p.26



### 1.2.2.1 Características sintácticas

#### Voz pasiva

En el lenguaje médico chino, se usan muchas expresiones pasivas para describir el estado de enfermo. Este rasgo que se aplica mucha voz pasiva se diferencia de la característica del lenguaje médico español. En el lenguaje médico español el abuso de voz pasiva es un problema, mientras que en el lenguaje médico chino es una característica y un requisito necesario en los documentos sanitarios. A lo largo del desarrollo del lenguaje médico, se ha establecido un consentimiento para cumplir en el trabajo médico. Se trata de un rasgo muy representativo. El motivo de este uso es reducir la redundancia, permitiendo que los médicos y los pacientes comprendan los síntomas y las claves de manera más intuitiva por medio de diluir el papel del sujeto de las oraciones.

*医学在长期的发展过程中早已形成鲜明的语言特点。如在医学诊断报告中常使用“未见……”、“未发现……”等较固定的语句表达式,来强调说明在此时此地提取的标本中该检验医生的一种主观断定[4]。在病案书写中表现为短句、无主句连贯使用的语言表达方式,如无尿、多汗、头痛乏力。(王茜, 2007)<sup>17</sup>*

En consonancia con Qian Wang (2007), el lenguaje médico chino ha desarrollado sus propias características diferentes de rasgos de otro idioma. Además, no se usan muchas oraciones largas con sudoraciones o de estructuras completas, sino se utilizan muchas expresiones muy cortas y concisas en la redacción o inscripciones de la enfermedad de los pacientes. Como, por ejemplo, si se redactan los registros, en caso de el enfermero no tiene dolor de cabeza, se describe “no visto el dolor de cabeza”, ignorando el sujeto, a fin de hacer los registros más precisos y fáciles de entender, así como mantener el uniforme de la estructura.

Esta característica es distinta de la del lenguaje médico español. Por un lado, este uso de voz pasiva no se aplica en oraciones, sino en locuciones con estructura “participio pasado +sustantivo”. Por otro lado, la mayoría de los casos en que se utiliza esta voz pasiva son registros o inscripciones del estado de enfermo. No se usa en todos tipos de documentos médicos, tales como consentimiento informado, receta, certificado, historia clínica, etc.

#### Expresión fija

En el lenguaje médico chino, se aplican con mucha frecuencia unas expresiones fijas para describir síntomas de enfermedad. Son relativamente fijas y requisitos básicos para seguir, además, es prohibido crear palabras o expresiones no existidas. La aparición de los problemas de mala combinación o incorrecta demostración no se debe a que los médicos carezcan de conocimientos profesionales, sino de dominación de las características del lenguaje médico.

---

<sup>17</sup> 王茜, (2007), 关于医院管理中医学语言质量管理的思考, 中国医学伦理学, 第20卷第4期 (Qian. W., (2007), Reflexiones sobre la gestión de calidad del lenguaje de la medicina tradicional china en la gestión hospitalaria, Ética médica china, volumen 20, número 4)

曾孔生教授还研究发现“病历的语言有相对固定的模型。常用“名词 Z”或“Z 名词”的固定形式来表述病情”。“Z”是固定的几个表示状况的词语，由医生根据患者实情选择，有一定的选择范围，“Z”前面或后面的名词是表示躯体部位和器官的专有名词，或表示人的生活状态、生命活动、人体机能的名词。如“发育”后面跟一个表示状况的“正常、不良、超常”，“营养”后面跟一个表示状况的“良好、中等、不良”，医生在书写病历时就只能在这三个词语中选一个，不能想当然的自造词语。从语用学上说这就是医学的语体特点。但在临床语言调查中，我们发现病历中常出现不符合医学语体特点的语言表述，如“排尿不能”、“发热面容”、“营养可，平静面容”、“角膜清晰”、“皮肤、黏膜未见黄染”。（王茜，2007）

De acuerdo con Wang (2007), hay modelos relativamente fijos para describir la enfermedad y los médicos no pueden elegir o crear palabras en los registros. Este modelo está compuesto por un sustantivo y un adjetivo: el sustantivo puede ser los órganos o partes del cuerpo, y el adjetivo está encaminada para describir el estado de los órganos o partes del cuerpo. Wang (2007) propone un ejemplo, que, para describir el estado del desarrollo de un niño, hay tres adjetivos para elegir “normal, malo, anormal”; para describir el estado de nutrición, solo hay tres opciones “bueno, medio, pobre”. Los médicos no tienen libertad de utilizar otras palabras adjetivas para hacer la descripción. Es importante conocer las expresiones fijas y aprender usarlas o traducirlas, es que, de esta manera, la traducción será más natural y de estándar. Si se traducen de manera informal documentos emitidos por el personal médico a pacientes, los pacientes pueden entender, pero falta de nivel profesional. No se puede ignorar las expresiones fijas especialmente en la traducción de los artículos o documentos solo enviados entre los expertos médicos.

### 1.2.2.2 Características morfológicas

#### Normalización

En el lenguaje médico chino, no se usan con frecuencia dialecto, jerga coloquial, expresión informal para indicar el diagnóstico. El médico suele describir los síntomas de pacientes con expresión normalizada. Es un principio básico y muy importante en los textos del ámbito sanitario. Por un lado, la ciencia médica es seria y requiere muy alto conocimiento específico y habilidad profesional. En cuanto a los documentos o registros escritos, hace falta presentar actitud rigurosa, es que tiene mucho que ver con la salud y la vida del paciente. Por otro lado, el uso de expresión normalizada puede reducir la posibilidad y aparición de malentendidos. No es posible que todos los personales médicos y pacientes provengan de una misma ciudad o un mismo país, lo que conduce a diferentes culturas y contextos entre sí.

比如诊断学中明确的要求在病历书写时“记录主诉要简明，应尽可能地用病人自己描述的症状，而不是医生对患者的诊断用语”。因为主诉主要是提供诊断线索，由于要“尽可能地用病人自己描述的症状”，但又不能用患者的言辞，语言特点是描述，而非解释，这就要求在进行医学书面语言表达时，医生必须对病人的语言进行转换，医生首先要把病人的口语、俚语转换成简明的医学规范用

语,如不用“拉肚子”而用“腹泻”或“稀水样便”。不用“喘不上气”或“掉气”,而用“气短”或“呼吸困难”。不用“肚子”,而用“腹部”。不用“厉害”而用“加剧”。不用“很少”而用“少许”。不用“第二天”而用“次日”。不用“打嗝儿”而用“暖气”。如患者自诉:近三天一会发冷,一会发热,用温度计量体温 38.6°C,胸脯右部疼。转换成规范的医学书面语应为“寒战、发热、右胸痛三天”。(王茜, 2007)<sup>18</sup>

En los registros médicos, los pacientes necesitan describir sus síntomas. Debido a la falta de conocimientos profesionales, los pacientes suelen expresar de manera muy coloquial. Sin embargo, los médicos tienen responsabilidad de transformar las palabras habladas en palabras escritas a fin de garantizar el estándar y exactitud del lenguaje médico. La normalización del lenguaje médico radica en una garantía de comprensión de todo el personal médico y personal sanitario. De todas maneras, es difícil comprender los dialectos o expresiones locales si no provienen de un mismo sitio. Como expone Wang (2007), en las inscripciones, se usan “diarrea” o “heces acusas sueltas” en lugar de “hacer cacas con demasiada frecuencia”; se usan “falta de aliento” o “dificultad de respirar” en vez de “pérdida de aliento”; se utiliza “vientre” en lugar de “barriga”; se aplica “agudización” en vez de “grave”; se usa “a nivel leve” en lugar de “poco”, etc. Ella también propone otro ejemplo: “si el paciente se describe: en los tres últimos días, me siento frío un rato, mientras que me siento caliente otro rato. He tomado la temperatura y la temperatura era 38.6°C. Además, también me duele la parte derecha del pecho. El lenguaje médico estandarizado debe ser: durante tres días, escalofríos, fiebre, dolor en el pecho derecho.” (Qian Wang, 2007)

## Prudencia

En el lenguaje médico chino, hacen falta mantener equivalencia y rigurosidad en todos los registros y evitar la aparición de ambigüedad y contradicción en la descripción. Para la ciencia médica, el conocimiento específico es una garantía de expresión exacta. Es que a veces, el significado de la oración no solo se limita a palabras, sino también contiene sentido complementario y añadido.

语句表达不严密,出现歧义。如“否认结核、肝炎等病”(应为:病史)。“注意观察有黑便、呕血等情况”。“舌尖、右侧舌缘可见 2 个溃疡面”(是一处一个,共 2 个,还是各处 2 个)。“病后精神 ...”(病后语义不清,是发病后还是病愈后?)(王茜, 2007)<sup>19</sup>

Traducción propia: Es fácil surgir la expresión no estricta y ambigua. Como, por ejemplo, “negar la tuberculosis, la hepatitis y otras enfermedades” (debería ser: historia médica), “prestar atención a la melena, vomitar sangre, etc.”, “verse 2

<sup>18</sup> 王茜, (2007), 关于医院管理中医学语言质量管理的思考, 中国医学伦理学, 第 20 卷第 4 期 (Qian. W., (2007), Reflexiones sobre la gestión de calidad del lenguaje de la medicina tradicional china en la gestión hospitalaria, Ética médica china, volumen 20, número 4), p87.

<sup>19</sup> 王茜, (2007), 关于医院管理中医学语言质量管理的思考, 中国医学伦理学, 第 20 卷第 4 期 (Qian. W., (2007), Reflexiones sobre la gestión de calidad del lenguaje de la medicina tradicional china en la gestión hospitalaria, Ética médica china, volumen 20, número 4), p87.

superficies de úlcera en la punta de la lengua y el margen lingual derecho” (¿cómo se distribuyen los dos?), “mental después de la enfermedad...” (¿después de su inicio o después de la recuperación?)” (Qian Wang, 2007)

En resumen, para un traductor sanitario o médico, conocer las características del lenguaje del idioma de salida y del idioma de llegada es imprescindible para llevar a cabo la traducción de manera exacta y excelente. Se trata de una de las competencias sobresalientes para ser un traductor calificado.

### **1.3 Características y dificultades de traducción sanitaria en España y China**

La traducción sanitaria se considera como una especialidad independiente por su complejidad en la conocimientos profesionales y habilidades específicas. Además, como es explica Ana Muñoz Miquel (2009), “es una actividad prescindible para el desarrollo de la comunicación y el conocimiento médico.” Sin embargo, a pesar de la dificultad para ser un traductor sanitario y el significado importante de traducción sanitaria, ahora según Ana Muñoz Miquel (2009) “se trata de un ámbito que apenas ha contado con el respaldo de educación universitaria y de la investigación.”<sup>20</sup> No existe un sistema maduro destinado a la formación de traductores médicos especializados. Tanto en España y en China está afrontando muchas dificultades la industria de la traducción sanitaria.

#### **1.3.1 Características de traducción sanitaria**

Como lo relatado de lo anterior, la traducción sanitaria es una traducción especializada, en que ha de comprender las características del lenguaje especializado médico. La medicina y el ámbito sanitario no solo es una ciencia académica, sino también una ciencia social.

#### **Amplitud de campo**

La medicinaa es una ciencia de campos muy amplios. “是处理生命的各种疾病或病变的一种学科，促进病患恢复健康的一种专业”（百度百科，2023）<sup>21</sup>. El fin consiste en tratar diversas enfermedades o lesiones para promover la recuperación de la salud. Es por esto que requiere las terminologías, conocimientos específicos de la traducción en este ámbito de acuerdo con el estado del paciente.

De acuerdo con el sitio web de Medine Puls, en España existen 4 campos generales: clínicas, quirúrgicas, medio-quirúrgicas y de laboratorio, incluyendo más de cuarenta especialidades médicas, como, por ejemplo: Alergia y Asma, Anestesiología, Cardiología, Dermatología, Endocrinología, Gastroenterología, Hematología, Nefrología, etc.<sup>22</sup>

En China, la manera de clasificar los campos de la ciencia médica es diferente. Los campos

---

<sup>20</sup> Muñoz-Miquel, Ana. “El perfil del traductor médico: diseño de un estudio de corte socioprofesional.” Panacea. X.30 (2009a): 157-167

<sup>21</sup> Enciclopedia Baidu, 2023

<sup>22</sup> Medine Puls: <https://medlineplus.gov/spanish/ency/article/001933.htm>, última citada: 2023

principales consisten en medicina forense, medicina tradicional china, medicina clínica, etc. Con respecto a la medicina básica, se abarcan: embriología humana, genética médica, inmunología, primeros auxilios, anatomía, ect. Por otro lado, en cuanto a la medicina clínica, se incluyen: Obstetricia, Ginecología, Pediatría, Oftalmología, Otorrinolaringología, Enfermedades Infecciosas, Dermatología, etc.<sup>23</sup>

Aquí se pone un sitio web para que pueda consultar si tiene interés: Nomenclatura de Ciencia y Tecnología de la UNESCO (SKOS) (<https://skos.um.es/unesco6/32/html>). Desde aquí nos da acceso a consultar nomenclatura de todas las ciencias y tecnologías y hacer análisis estadístico, lo que está destinado a facilitar la investigación y estudio.

En resumen, en cada campo, cada especialidad se engloba diferentes terminologías. Se trata de un punto difícil para un traductor. De acuerdo con Albir (2001), “los principales problemas que se encuentran en la traducción consisten en: problemas lingüísticos, que son discrepancias entre los dos idiomas en diferentes aspectos léxicos, morfosintácticos, textuales, estilísticos; problemas pragmáticos; problemas extralingüísticos, que consisten en cuestiones temáticas, estilísticos y textuales; y problemas instrumentales.”<sup>24</sup>

Entonces necesitamos prestar atención a la amplitud y completitud en este ámbito. Para un traductor, dominar la terminología, estilo, tema, conocimiento, herramienta adecuada, etc. de un campo es la clave para llevar a cabo la traducción. Si no conociera esta característica, podría causar errores en la traducción, e incluso conducir a resultados graves para las dos partes.

### **Situaciones comunicativas y géneros textuales**

En las escenas sanitarias, los usuarios que acuden al ámbito sanitario provienen de diferentes categorías, con diferentes intenciones. Puede ser investigadores que tienen necesidad de hacer investigación sanitaria en centro médico. También puede ser los personales sanitarios que van a hacer una conferencia académica de intercambio. Teniendo en cuenta en diferentes participantes, diferentes sitios, las situaciones de comunicación varían mucho. Cabe destacar que, debido a diferentes niveles de conocimientos profesionales de participantes, el nivel de dificultad de los documentos para traducir tiene gran diferencia. Es imposible que el nivel de dificultad de traducir prescripciones sea igual que traducir un artículo científico de medicina. Como se explica Ana Muñoz Miquel (2016), “La variedad de situaciones de comunicación implica las participantes de interlocutores de perfiles y conocimientos diversos (investigadores, profesionales de la salud, pacientes, público general, etc.), lo que influye en el grado de especialización y de formalidad de la comunicación.”<sup>25</sup>

En cuanto a las situaciones de comunicación y los géneros textuales, según Miquel (2016), se

---

<sup>23</sup> Enciclopedia Baidu, 2023

<sup>24</sup> Albir, A. H. (2001). Traducción y traductología / Translation and translation studies: Introducción a la traductología / Introduction to Translation. Catedra Ediciones.

<sup>25</sup> Muñoz Miquel, Ana, (2016), La traducción médica como especialidad académica: Algunos rasgos definitorios. Hermeneus. Revista de la Facultad de Traducción e Interpretación de Soria Núm. 18 Pág. 235-267

puede clasificar de la siguiente manera: farmacéutico, empresas de traducción, investigación, editorial, instituciones europeas, y salud pública y atención sanitaria.

**Tabla 6**

*Participantes, socio-funciones, sectores y tipos textuales en las comunicaciones médicas (Citado de Ye Lu, 2020, Montalt y González Davies, 2014, pp. 62-70; Mayor Serrano et al., 2004, pp. 130-132)<sup>26</sup>*

<p><b>Participantes</b>          investigadores          personal sanitario          farmacéuticos          instituciones/organismos de salud          estudiantes de medicina          pacientes          público general</p>	<p><b>Socio-funciones</b>          difusión de investigación médicas          comunicación en el centro de salud          comunicación con los pacientes          educación del público          enseñanza a los estudiantes de medicina          aprobación de nuevos medicamentos          regulación de productos sanitarios</p>
<p><b>Sectores</b>          centro de investigación          hospital o otro centro de salud          laboratorio          industria farmacéutica          instituciones/organismos de salud          editoriales de ciencia médica</p>	<p><b>Tipos textuales</b>          artículos científicos de medicina          prospectos de medicamentos          consentimientos informados          prescripciones          informes clínicos          folletos de información a pacientes          campañas institucionales          materiales pedagógicos de medicina          libros/maunales no especializados          artículos divulgativos          documentos en televisión sobre salud</p>

Además, las diversas comunicaciones requieren los géneros textuales correspondientes. De acuerdo con Miquel (2016), en cuanto a los documentos médicos o exclusivos del ámbito, se abarcan el folleto de información para pacientes, el manual para promover la salud pública, el prospecto de medicamento, el protocolo de ensayo clínico, artículos divulgativos, materiales pedagógicos de medicina, el vademécum, el cuestionario de salud, el consentimiento informado, la guía clínica, el resumen para pacientes, los resultados de diferentes pruebas y exámenes, etc.<sup>27</sup> La comprensión de los géneros textuales puede facilitar la preparación de la traducción. Después de saber los participantes, los géneros textuales, el traductor puede buscar textos paralelos pertinentes, lo que da oportunidad de mejorar la calidad de traducción y evitar errores.

<sup>26</sup> Ye Lu, la terminología aplicada a la traducción médica: elaboración de un glosario bilingüe chino-español, TFM, 2020, Máster Universitario en Comunicación Intercultural, Interpretación y Traducción en los Servicios Públicos

<sup>27</sup> Muñoz Miquel, Ana, (2016), La traducción médica como especialidad académica: Algunos rasgos definitorios. Hermeneus. Revista de la Facultad de Traducción e Interpretación de Soria Núm. 18 Pág. 235-267

## Terminología y culturas

Como se argumenta lo anterior, el lenguaje especializado tiene su propia terminología en consonancia con el campo específico. Y es una zona desconocida para uno que no ha aprendido este ámbito. Para la traducción médico, el papel de terminología destaca más. Al mismo tiempo, según M. Teresa Cabé Castellví (2004), la terminología como disciplina está basada en conocimiento especializado para facilitar la comunicación especializada.<sup>28</sup> En otras palabras, en la traducción general, como la conversación diaria, sin dominio de terminología, se puede realizar la traducción de los documentos sin obstáculos. No obstante, en el ámbito sanitario, es imposible comprender el texto sin preparación de glosario o terminología de antemano.

Como, por ejemplo, una frase muy simple, “el enfermo tiene síntomas de diarreas, faringitis, amigdalitis, otitis y sinusitis”, e incluso en esta frase no hay gramática compleja, solo es una frase simple con estructura “sujeto+verbo+objeto”. Sin conocer los términos “diarrea”, “faringitis”, “amigdalitis”, “otitis”, “sinusitis”, es imposible llevar a cabo la traducción de esta frase. Como menciona M. Teresa Cabé Castellví (2004), “no existe materia especializada alguna sin terminología”, se trata de una característica más destacable en el lenguaje especializado.

Por un lado, solo conocer las terminologías no es suficiente. De acuerdo con Miquel (2016), “la imprecisión en el uso de la terminología puede afectar de manera directa a la comprensión y dar lugar a interpretaciones erróneas también en el texto meta.”<sup>29</sup> Según lo argumentado en el apartado anterior, el uso de “falsos amigos” es un problema en este sentido. Por ejemplo, “condition of patients” no se puede traducir como “condición de pacientes”, sino “estado fiscal de pacientes”. Por consiguiente, para un traductor sanitario, hay muchos aspectos en que deben tener en cuenta. Solo por medio de comprender significado de terminologías, se puede provocar malentendidos o confusión, por lo que saber cómo utilizar los términos, saber cómo adaptarlos, son más importantes.

Por otro lado, la aplicación de siglas y acrónimos consiste en un punto difícil para la traducción médica. Se usan con mucha frecuencia en los resultados de exámenes con el fin de mostrar de manera más concisa y visual los indicadores de diversos valores. Entonces, aprender utilizar diccionario y recursos de herramientas también cuenta mucho para los traductores médicos. De acuerdo con Javier Yetano Laguna y Vicent Alberola Cuñat (2011), “la formación de los términos interviene raíces griegas o latinas, y se exponen siglas médicas, otras siglas, epónimos y términos médicos relacionados con la codificación de las altas hospitalarias.”<sup>30</sup>

## Figura 1

---

<sup>29</sup> Muñoz Miquel, Ana, (2016), La traducción médica como especialidad académica: Algunos rasgos definitorios. *Hermeneus. Revista de la Facultad de Traducción e Interpretación de Soria* Núm. 18 Pág. 256

<sup>30</sup> Laguna, J. Y., & Cuñat, V. A. (2011). *Diccionario de siglas médicas*. SEDOM.

siglas de indicadores de detección bioquímica (复旦大学实验动物科学部(Departamento de Ciencia de Animales de Laboratorio, Universidad de Fudan, 2023)<sup>31</sup>

生化检测指标				
一. 肝功能检查项目		四. 肾功能检查项目		
1	碱性磷酸酶	ALP_2c	15 肌酐	
2	γ-谷氨酰转氨酶	GGT	16 尿酸	
3	丙氨酸氨基转移酶	ALT	17 尿素氮	
4	天门冬氨酸氨基转移酶	AST	五. 心血管疾病检测项目	
5	总蛋白	TP	18 肌酸激酶	CK_L
6	白蛋白	ALBP	19 乳酸脱氢酶	LDLP
7	总胆红素	TBIL_2	六. 微量元素	
8	直接胆红素	DBIL_2	20 钙离子	CA
9	胆碱酯酶	CHE	21 镁离子	MG
二. 血脂检测项目		22 无机磷	IP	
10	高密度脂蛋白胆固醇	D-HDL	23 铁离子	IRON_2
11	低密度脂蛋白胆固醇	LDL	七. 电解质	
12	甘油三酯	TRIG_2	24 钠	Na
13	总胆固醇	CHOL_2	25 钾	K
三. 糖疾病检测项目		26 氯	Cl	
14	葡萄糖	GLUH_3		

Según Miquel (2016), “las monografías dedicadas a la traducción médica no suelen abordar ls problemas culturales como un rasgo característicos...”<sup>32</sup> Surge necesidad de destacar el factor cultural en la traducción médica de español a chino, porque la medicina tradicional china engloban muchos términos filosóficos y conceptuales de conciencia.

## Figura 2

Diagrama general de la teoría de la medicina china (Enciclopedia Baidu, 2023)



En la Figura 2, se describen la correspondencia entre los cinco elementos, es decir, el oro, la madera, el agua, el fuego y la tierra, y los distintos sistemas del cuerpo y las interacciones entre los elementos. La tierra corresponde al sistema digestivo, y su órgano correspondiente es el bazo; el fuego corresponde al sistema endocrino y su órgano correspondiente es el corazón; el oro, del sistema respiratorio, el pulmón; la madera, al sistema inmunitario, el hígado; el agua, al sistema circulatorio, el riñón. La tierra favorece al oro, el oro favorece al agua; el agua favorece a la madera; la madera favorece al fuego; el fuego favorece a la tierra. A la inversa, la tierra desfavorece al agua; el agua desfavorece al fuego; el fuego desfavorece a oro; el oro desfavorece a la madera; la madera desfavorece a la tierra. Se trata de una relación de

<sup>31</sup> 复旦大学实验动物科学部(Departamento de Ciencia de Animales de Laboratorio, Universidad de Fudan): <http://labanimal.fudan.edu.cn/26586/list.htm>, última citada: 2023

<sup>32</sup> Muñoz Miquel, Ana, (2016), La traducción médica como especialidad académica: Algunos rasgos definitorios. Hermeneus. Revista de la Facultad de Traducción e Interpretación de Soria Núm. 18 Pág. 235-267



equilibrio cuyo concepto central es la circulación. Como explica Eric Marié (1998), “en la medicina china se considera que ningún fenómeno vital puede ser analizado fuera de su contexto orgánico y universal.”<sup>33</sup>

Es muy difícil entender los conceptos en la medicina china, pero es cierto que es un componente muy importante para el lenguaje médico chino. Así, para un traductor dedicado a esta dirección, estudiar las características culturales es imprescindible.

### **1.3.2 Dificultades de traducción sanitaria**

#### **1.3.2.1 Falta de formación sistemática**

Se proponen las características de traducción sanitaria en el apartado anterior, entonces, sabemos que para traducir un texto documental hay muchos aspectos para preparar: conocer y saber el uso de términos de diferentes campos de medicina, comprender contextos culturales, dominar conocimientos pertinentes, etc. Entonces, para ser un traductor capacitado, hace falta recibir la educación y formación profesional a fin de garantizar la calidad de textos traducidos. Según Franz Pöchhacker (2004), “por medio de una formación especializada se puede adquirir conocimientos necesarios, y esto es una condición para que una ocupación sea reconocida como profesión.”<sup>34</sup> En España existe una selección y examen llamado Traductor Jurado, pero en el ámbito sanitario no existe un examen reconocido para obtener un título de traductor médico, lo que significa que todavía no hay una norma uniforme para calificar la capacidad y competencia de traducción médica.

La falta de formación sistemática trae consigo una serie de problemas. Sin formación profesional, no se puede garantizar la calidad de la traducción médica de los traductores, posiblemente conduciendo a la pérdida de confianza de los clientes. Sin reconocimiento de los clientes, el mercado laboral no va a prestar tanta atención al traductor médico, entonces, es difícil convertir en una profesión, un trabajo profesional. En consecuencia, no puede recibir la garantía de protegido de los reglamentos de los gobiernos.

Actualmente, a nivel mundial existen unos cursos para los traductores en el ámbito sanitario. Como se expone Abril Martí (2006), hay Sudáfrica, Community Health Interpreter Training and Employment Programme, Western Cape, en Estados Unidos, Bridging the Gap, programa de formación básica/intermedia de Cross Cultural Health Care Programme (CCHCP), Seattle, Washington, Programa de salud pública de Massachusetts, formación estándar para personal bilingüe y voluntarios de centros sanitarios, Massachusetts, en Austria, Curso de Interpretación para Personal Sanitario Bilingüe, de la Cooperación de Hospitales Municipales de Viena, etc.<sup>35</sup>

---

<sup>33</sup> Marié, E. (1998). Compendio de medicina china: fundamentos, teoría y práctica (Vol. 180). Edaf.

<sup>34</sup> Pöchhacker, F. (2004). *Introducing Interpreting Studies*. Londres/Nueva York: Routledge

<sup>35</sup> Abril Martí, M.I. (2006). *La interpretación en los servicios públicos: Caracterización como género, contextualización y modelos de formación*. Hacia unas bases para diseño curricular. Universidad de Granada

De acuerdo con Martínez de Cañas Pagola, Leire Cristina (2017)<sup>36</sup>, existen los siguientes cursos para la traducción médica: Curso de especialista en traducción médica EN-ES, de Ágora Formación, Seminario de traducción médico-farmacéutica, de Estudio Sempere, Samario Problemas, métodos y cuestiones candentes en traducción médica, de la Universidad Internacional Menéndez Pelayo, Taller de traducción biosanitaria: Análisis y reescritura del texto, de Con Trazo Firme, Prácticum de traducción médica, de Aula SIC, Traducción de medicina oriental (EN-ES), de Aula SIC, Psicopatología: Teoría y práctica para traductores, de Aula SIC, etc.

En China, existe una prueba profesional para elegir traductores, pero no se limita a un ámbito, sino que abarca un tema general. La Prueba de Acreditación de China para Traductores e Intérpretes (CATTI) es un examen de calificación vocacional a nivel estatal, que está diseñado para evaluar la competencia de los candidatos. Aunque no está enfocado en la traducción médica, consiste en una prueba de traducción entre español-chino. Esta prueba está encaminada a diferentes pares de idiomas estableciendo una norma uniforme de capacidades y conocimientos de traductores.

Al mismo tiempo, afortunadamente ya hay varias universidades que establecieron cursos para dictar cursos de TISP (Traducción e interpretación en los servicios públicos), como la Universidad de Alcalá, la Universidad de Granada, la Universidad de Murcia, la Universitat Jaume I, La Universidad del País Vasco, la Universidad Alfonso X el Sabio, la Universidad Autónoma de Barcelona, etc. Además, también se han fundado asociaciones específicas para que los traductores e intérpretes tengan acceso a los recursos y compartir conocimientos y experiencias. Como por ejemplo, en cuanto a las asociaciones de traductores e intérpretes, se incluyen Asia-Pacific Association for Machine Translation (AAMT), Asociación Europea de Terminología (AET), Asociación Internacional de Traductores e Intérpretes (AI-TI), International Medical Interpreters Association (IMIA), Translators Association of China (TAC), Asociación Profesional de Traductores e Intérpretes Judiciales y Jurados (APTIJ), Asociación de Formadores, Investigadores y Profesionales de la Traducción e Interpretación en los Servicios públicos (AFPTISP), etc.

Desde mi punto de vista, aunque “la elaboración de dicho programa fue muy costosa y minuciosa y requirió horas incontables de trabajo y esfuerzo” (Garcés, Carmen Valero, and Mustapha Taïbi, 2006), hace falta establecer programas de formación y cooperar con las universidades para que desempeñen su iniciativa en promoverla.

### **1.3.2.2 Falta de código deontológico**

Antes que nada, un código de conducta, o sea código ético o código deontológico, consiste en un conjunto de principios para regular el comportamiento de los trabajadores. En cuanto a cada profesión, seguir un código ético es una garantía para el cumplimiento del trabajo, incluido para los traductores e intérpretes en el ámbito sanitario. En cambio, la verdad es que

---

<sup>36</sup> Martínez de Cañas Pagola, L. C. (2017). Análisis de la Formación en Traducción médico-sanitaria en España: una primera aproximación. Universitat Jaume I

en España y China todavía no existe un código deontológico establecido para los traductores que trabajan en el ámbito sanitario.

Como expone Augusto Hortal Alonso (2002), la deontología consiste en normas y principios escritos.<sup>37</sup> Igualmente, aplicados en la traducción médica, los códigos éticos están destinados a dar instrucciones para los traductores. Solo con una serie de códigos deontológicos establecidos para la traducción médica, se conduce una norma para guiar la dirección de traducción cómo llegar al nivel bueno.

Generalmente, los traductores deben seguir las siguientes normas básicas: imparcialidad, neutralidad, exactitud, confidencialidad y integridad. Pero es una serie de código para todos los traductores, de carácter más general. No se puede adaptar a la traducción médica cien por cien. Cuando aparece el dilema ético, de acuerdo con Hortal Alonso (2002), el traductor debe tomar medidas siguiendo los siguientes pasos: 1) conocer las situaciones y circunstancias; 2) percibir las posibilidades y los diferentes cursos de acción; 3) distinguir el significado de lo que está en juego y las semejanzas y diferencias con otras situaciones; 4) aplicar los conocimientos profesionales a fin de clasificar la situación y los cursos de acción; 5) identificar a qué principios apela la situación.<sup>38</sup>

En resumen, en caso de dificultades o problemas, de primero desentraña este problema y anticipa las posibles consecuencias, aplica la habilidad para buscar la solución. En otras palabras, necesita tomar iniciativa para ajustar la traducción. Según Lázaro Gutiérrez (2009), “el traductor no es un autómatas, se humaniza, sino puede asumir la responsabilidad para mantener la comunicación, eliminando las desigualdades, malentendidos entre dos partes de interlocutores.”<sup>39</sup>

Desde el punto de vista de Valero Garcés (2014), el código ético para los traductores e intérpretes sanitarios está basado en los siguientes cuatro principios:

1. Integridad: El traductor e intérprete debe juzgar el límite de sus propias capacidades, si no está capacitado, rechazará trabajos.
2. Confidencialidad: El traductor e intérprete tiene obligación de proteger la intimidad y todo lo que interviene en la comunicación.
3. Imparcialidad: El traductor e intérprete debe evitar ponerse del lado de ninguna parte de la conversación.
4. Precisión/Fidelidad: El traductor e intérprete debe respetar el texto original, transmitiendo un mensaje fiel, sin añadir ni omitir datos o informaciones.<sup>40</sup>

En China, en cuanto a códigos éticos, hay una norma muy reconocida, tres principios: 信, 达, 雅. De acuerdo con la Enciclopedia Baidu, 信: requiere que la traducción no contradiga

---

<sup>37</sup>Hortal Alonso, Augusto (2002): *Ética general de las profesiones*. Bilbao: Desclée de Brouwer.

<sup>38</sup> Hortal Alonso, Augusto (2002): *Ética general de las profesiones*. Bilbao: Desclée de Brouwer.

<sup>39</sup> Lázaro Gutiérrez, R. (2009). Dilemas éticos del traductor/intérprete en los servicios públicos. *Translation, Interpreting, Social Activism*, sin paginar.

<sup>40</sup> Valero-Garcés, C. et al. (Ed.). (2014). (Re)considerando ética e ideología en situaciones de conflicto/ (Re)visiting ethics and ideology in situations of conflict. Alcalá de Henares: Servicios de publicaciones de la Universidad de Alcalá

el texto original, es decir, la traducción debe ser precisa, sin desviaciones, sin omisiones y sin añadir o quitar arbitrariamente. *达*: requiere que la traducción no esté atada a la forma del texto original, debe ser fluida y clara. *雅*: requiere que las palabras utilizadas en la traducción deben ser adecuados y el texto traducido debe ser elegante y sencillo. Pero, estos tres principios, en mi opinión, se pueden aplicar en la traducción literaria sino en la traducción médica. Es que el lenguaje sanitario no tiene diversas formas de expresión artística, ni necesidad de expresión elegante.

*“信”“达”“雅”它是由我国清末新兴启蒙思想家严复提出的，他在《天演论》中的“译例言”讲到：“译事三难：信、达、雅。求其信，已大难矣！顾信矣，不达，虽译，犹不译也，则达尚焉。”（百度百科，2023）<sup>41</sup>*

En conclusión, el traductor e intérprete sanitario todavía no se admite como una profesión específica. La mayoría de los trabajadores son traductores e intérpretes generales quienes tienen conocimientos pertinentes o personales médicos quienes dominan en algún sentido los idiomas, por lo que todavía faltan el reconocimiento de la sociedad y el protegido legal. Es por esto que la traducción e interpretación sanitaria está enfrentándose con muchos obstáculos en el proceso de avance. Si se quiere amplificar su impacto y formalizarlo, para China y España, a nivel nacional, hace falta prestar atención a establecer un sistema de formación y una serie de códigos deontológicos. Para realizar este objetivo, es necesario la colaboración de diferentes categorías: sociedad, gobierno, universidades, institutos, asociaciones, traductores, intérpretes, etc.

## **2. La traducción automática en contextos especializados**

### **2.1 Traducción automática: definición y tipología de sistemas**

#### **2.1.1 Definición de TA**

Antes que todo, la traducción automática es un tipo de traducción realizada por la tecnología o máquina sin intervención humana. La traducción es una acción para transferir mensaje en un idioma al en otro idioma. La TA también está destinada a este objetivo, mientras que la manera de traducir no depende de la humana sino de la tecnología. Es un producto del desarrollo de la tecnología avanzada y un producto ante la necesidad de análisis de datos.

Se exponen diferentes opiniones en cuanto a las definiciones de TA:

La TA se concibe como la aplicación de la tecnología informática a la traducción de textos de una lengua a otra sin intervención humana. (María del Mar Sánchez Ramos, 2020)<sup>42</sup>

---

<sup>41</sup> Enciclopedia Baidu, 2023

<sup>42</sup> Sánchez Ramos, et al. Traducción automática: conceptos clave, procesos de evaluación y técnicas de posesición. Comares, 2020.

La traducción automática es traducción automatizada. Es el proceso mediante el cual se utiliza software de computadora para traducir un texto de un lenguaje natural a otro. (SYSTRAN, 2023)<sup>43</sup>

*机器翻译，又称为自动翻译，是利用计算机将一种自然语言（源语言）转换为另一种自然语言（目标语言）的过程。它是计算语言学的一个分支，是人工智能的终极目标之一，具有重要的科学研究价值。（百度百科，2023）<sup>44</sup>*

Traducción propia:

La traducción automática, también conocida como traducción de máquina, es el proceso de convertir una lengua de partida en otra lengua de llegada mediante un ordenador. Es una rama de la lingüística computacional, uno de los objetivos últimos de la inteligencia artificial, y tiene un gran valor para la investigación científica. (Enciclopedia Baidu, 2023)

En la Figura 3, se enumeran las herramientas de traducción automática a nivel mundial. En cuanto a plataformas generales, existe estas herramientas: Alibaba Translate, Amazon, Aperlum, AISA, Baidu, Google Translate, GTCOM, IBM Watson, Language Weaver, Lingmo, LingvaNex, Microsoft, Mirai, ModernMT, Omnicen Lenguaje Estudio, Papage Naver, Promt, Sogou, Systran, Rozette T-400, Tencent Cloud, TranslateMe, Translatewise, XI8.ai, Yandex Translate. En cuanto a Mi TOOLKITS, hay las siguientes herramientas: Farirseq, Marian C++, ModernMT, Moses, Nematus, OpenNMT, seq2seq, sockeye, THUMT, Yandex RNNLM. Al respecto de Niche/legacy, Amebis Presis, AnyLangTech, Aootel, Babylon, Croalanguage, Dayi, Kingsoft, Leo Translate, Linguateo, LogoVista, OpenTrad, QCRI, Reverie, Reverso, Salesforce, SAP, Secure MT, Sunda MT, Textra NICT, TextShuttle, Wordlingo, Omniscien, Verto Lingua Custodia. Y atendio Custom/trainale, CloudTranslation, Globalese, Iconic, KantanMT, Lilt, Lucy LT, Newtranx, Niutrans, Omniscien, PangeaMT, Prompsit, Promt, Language Weaver, SmartMATE, Systran, Tide, Tarjama Advance MT, ModernMT, GlobaLink AL/MT, Lenguaje Wire Translate.

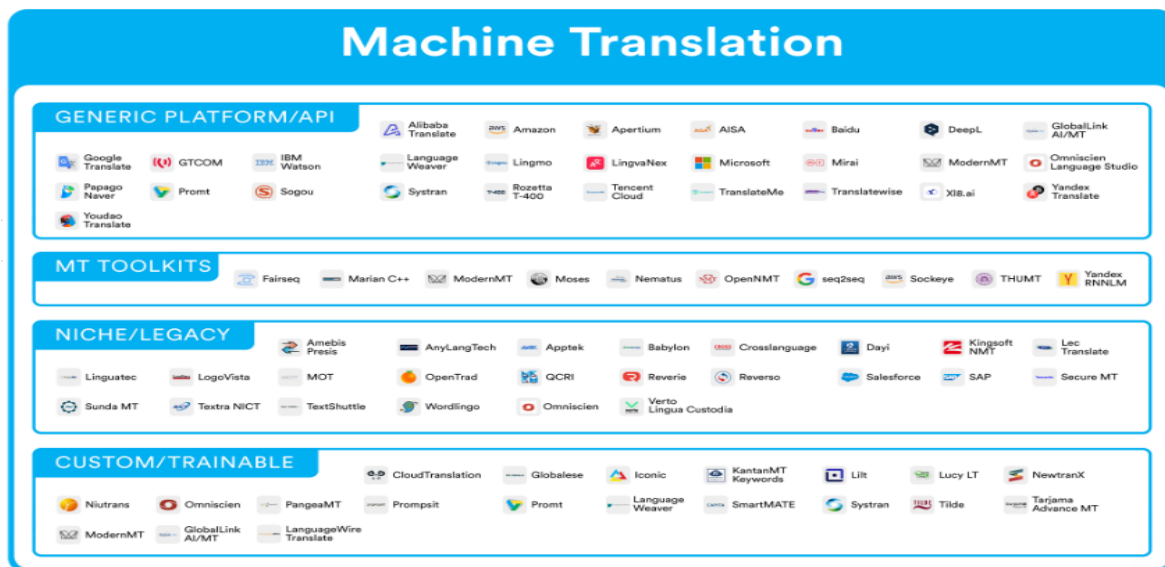
### **Figura 3**

*LANGUAGE TECHNOLOGY LANDSCAPE (Nimdzi Lenguaje Tecnología Altas 2021)*

---

<sup>43</sup> SYSTRAN: <https://www.systransoft.com/es/systran/tecnologia/que-es-la-traduccion-automatica/>, última citada: 2023

<sup>44</sup> Enciclopedia Baidu, 2023



Excepto de las herramientas enumeradas, en China hay otros instrumentos: traductor iFLYTEK, Youdao Translate, CaiYun Translate, Bi Ying Translate, Haici Translate, Hujiang Xioa D, y CNKI Translate. Los motores de traducción tienen sus propias bases de datos, terminologías y corpus, por lo que la calidad y las características de sus traducciones son variadas.

### 2.1.2 Desarrollo de la TA

Según lo señalado de Sánchez (2020), la traducción automática es un “fruto de la investigación y el esfuerzo de científicos”.<sup>45</sup> Así, la aparición de la TA nació con el desarrollo tecnológico y la industrialización. Además, los hitos e hilos de desarrollo tienen que ver con los puntos claves de creación de tecnología.

En cuanto al desarrollo de la TA, se resumen las fases significativas en base a la teoría de Pilar Hernández Mercedes (2002)<sup>46</sup>, y Yu Qingqing, Shi Fengjun (2012)<sup>47</sup>,

-**Embrión** de diccionario mecánico en el siglo XVII.

-**Planificación** de la TA a los años 30 del siglo XXV: Concretamente, en 1933, surgían dos patentes para “máquinas de traducir”, una en Francia, y otra en Rusia. Y en 1936, nació la primera computadora electrónica moderna del mundo, ENIAC. Después, Warren Weaver propuso la idea de traducir por medio de computadora y publicó un informe en julio de 49 en que se planeó la idea de TA.

-**Fase de pionero:** En 1954, se realizó el primer experimento público de un proyecto de traducción automática, con la cooperación de Universidad de Georgetown y IBM. Y después, el

<sup>45</sup> Sánchez Ramos, et al. Traducción automática: conceptos clave, procesos de evaluación y técnicas de posesición. Comares, 2020.

<sup>46</sup> Mercedes, P. H. (2002). En torno a la traducción automática. Cervantes, 1(2), 101-117.

<sup>47</sup> 郁青青, 石冯珺, (2012), 机器翻译系统的历史与研究现状分析, A, 淮海工学院学报(人文社会科学版), 第10卷第12期 (Qingqing Yu, Jun Shi, (2012), Análisis de la historia y el estado de la investigación de los sistemas de traducción automática, A, Journal of Huaihai Institute of Technology (Humanities and Social Sciences Edition), Vol. 10, No. 12

estudio en la TA mantenía activo, es que, en el contexto de Guerra Frío, las dos grandes potencias de los Estados Unidos y la Unión Soviética han invertido mucho dinero en TA por necesidades militares. Centrándose en la traducción de emparejamiento de idiomas inglés y ruso, el objeto principal de la traducción son los documentos científicos y técnico. Durante la época (1955-1965), se creyeron los primeros sistemas, conocidos como “sistemas de primera generación”. En 1957, Chomsky planeó un modelo formal de descripción lingüística para llevar al cabo un análisis sintáctico y semántico de las lenguas de origen y de destino. En 1964, la Academia Nacional de los EEUU creó ALPAC (Automatic Language Processing Advisory Committee).

**-Desarrollo en suspenso:** en 1966, ALPAC publicó un informe en que se reveló la baja eficiencia, menos exactitud de la TA, al mismo tiempo se recomienda parar de proporcionar las subvenciones, es que en cuanto a la TA, no existiría expectativa previsible y futuro

**-Fase de reanimación:** Con el desarrollo de la informática, la lingüística, y la demanda de servicios de información social, la TA comenzó a recuperarse. Se han desarrollado diversos sistemas de traducción, como Weinder, URPOTRAA, TAUM-METEO, etc. Cabe destacar los sistemas SYSTRAN, LOGOS y METAL.

**-Fase de desarrollo:** Desde los años 90, se han dado importantes pasos en la investigación de la TA, como la creación de *estaciones de trabajo* para traductores profesionales. En 1993, se propuso un modelo de traducción basado en la alineación de palabras, lo que consiste en un símbolo de nacimiento de los métodos modernos de traducción automática. En 2003, se propuso un modelo logarítmico lineal y su método de entrenamiento con pesas.

**-Fase de gloria:** En 2006, se lanzó al mercado Google Translate, contribuyendo una ola de investigación estadística sobre TA. En 2003, se propuso una nueva arquitectura de codificador-decodificador de extremo a extremo para la traducción automática. El modelo utiliza una red neuronal convolucional (CNN) para codificar un texto de origen en un vector continuo y después utiliza una red neuronal recurrente (RNN) como decodificador para convertir el vector de estado en el idioma de destino. Se pueden considerar como el comienzo de la traducción automática neuronal (NMT). En 2014, se propuso Sequence to sequence learning with neural networks, que está basado en la arquitectura codificador-decodificador, donde el codificador y el decodificador son de estructuras RNN, usando LSTM. En 2015, se agregó el peso de atención al marco codificador-decodificador. En 2017, se abandonaron la CNN y la RNN tradicionales, y toda la estructura de la red está compuesta completamente por el mecanismo de atención. Las principales innovaciones de Transformar son los cabezales múltiples y la autoatención.

### 2.1.3 Tipología de sistemas

De acuerdo con el apartado anterior, durante el desarrollo de la TA, aparece de manera continua diferentes sistemas, lo que se produce debido a la investigación de los ingenieros y científicos. Como expone Sánchez (2020), los sistemas de TA se pueden dividir en los siguientes tipos:

- ✓ Sistemas de traducción automática basados en reglas
- ✓ Sistemas de traducción automática basados en estadística

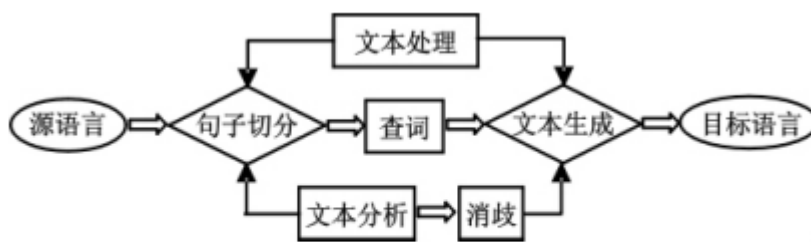
- ✓ Sistemas de traducción automática basados en ejemplos
- ✓ Sistemas de traducción automática híbridos
- ✓ Sistemas de traducción automática basados en redes neuronales

### Sistemas de traducción automática basados en reglas

En primer lugar, al respecto al sistema de TA basada en reglas, según Yuan X. Y. (2011), “经典的基于规则的机器翻译技术主要包括源语言的文本处理、词典构成、文本分析, 源语言到目标语言的词典查询翻译, 目标语言文本生成等技术”<sup>48</sup>, es decir, “las tecnologías generalmente consisten en procesamiento de textos del idioma de origen, composición del diccionario, análisis de textos, traducción de consultas de diccionario del idioma de origen al idioma de destino, así como generación de texto del idioma de destino, etc.” Según la figura 4, se explica el proceso de traducción de este sistema. Después de ingresar el idioma de origen, el sistema segmenta la oración, realiza consulta de vocabulario, analiza el texto para seleccionar palabras y luego realiza un procesamiento de texto para generar texto en el idioma de destino.

**Figura 4**

*Ilustración del sistema de traducción clásico (Yuan Xiaoyu, 2011)*



**图 1 经典机器翻译系统图示**

Como propone Yuan (2011), la dificultad del sistema de TA basado en reglas consiste en tres aspectos. En primer lugar, como análisis de texto es un proceso necesario para la traducción, TA necesita seleccionar una palabra más exacta dentro de muchos sinónimos. Es difícil para el sistema en regla, es que la combinación de las palabras sinónimas contiene diferentes significados. Este punto puede causar la traducción falsa o incorrecta por las palabras seleccionadas. En segundo lugar, desde la estructura del idioma de origen y de destino, la diferencia de la estructura lleva dificultad a TA en la segmentación de oraciones. La segmentación ejerce efecto de manera directamente a la relación entre los grupos de significado en una oración. Si no se segmenta muy adecuadamente, a la traducción falta la lógica. Por último, la expresión de emociones y sensaciones consiste en un problema permanente y un punto que atrae mucha discusión del público. En el texto del idioma de origen se contiene la sensación de tristeza, alegría, preocupación, y es muy difícil para TA expresarla solo por medio de la combinación de palabras.

<sup>48</sup> 袁小于.(2011).基于规则的机器翻译技术综述. 重庆文理学院学报(自然科学版)(03),56-59.doi:10.15998/j.cnki.issn1673-8012.2011.03.016.



Sin embargo, es cierto que los sistemas de TA basados en reglas tienen sus ventajas, según propone Sánchez (2020)<sup>49</sup>, se resumen los dos aspectos beneficiosos: por un lado, como está basado en reglas, se puede perfeccionar o cambiar las reglas para corregir problemas o errores. Se puede considerar como una orden de código de máquina o una fórmula. Tiene capacidad de controlar la lógica básica, así es fácil de ajustarlo, aunque a veces la generación del idioma de origen no es suficientemente flexible. Por otro lado, en caso de buena preparación de glosario y reglas un ámbito muy específico, la TA puede generar traducción de alta calidad. Y una de las herramientas de este sistema es Apertium.

### Sistemas de traducción automática basados en estadística

Antes que nada, según la figura 5, el proceso de traducción de este tipo de sistema consiste en que el texto del idioma de destino pasa por el canal de ruido, donde el sistema calcula y busca, y después encuentra la traducción más adecuada del idioma de destino a través de la decodificación.

可见，统计机器翻译系统的任务就是在所有可能的目标语言（翻译意义上的目标语言，也就是噪声信道模型意义上的源语言）的句子中寻找概率最大的那个句子作为翻译结果。其概率值可以使用 Bayes 公式得到（公式中的 T 是在翻译意义上目标语言，S 是在翻译意义上的源语言）：

$$P(T|S) = \frac{P(T) P(S|T)}{P(S)} \quad (\text{Feng Zhi Wei, 2010})^{50}$$

Traducción propia: De acuerdo con lo relatado anterior, la tarea del sistema de TA basado en estadística consiste en encontrar la oración con la mayor probabilidad entre todas las oraciones posibles en el idioma de destino (el idioma de destino en el sentido de traducción, es decir, idioma de origen en el sentido del modelo de canal de ruido) como resultado de la traducción. El valor de probabilidad se puede obtener utilizando la fórmula de Bayes (T en el sentido de traducción es el idioma de destino y S es el idioma de origen en el sentido traducido).

### Figura 5

*Simulación de canales de ruido (Feng Zhiwei, 2010)*

<sup>49</sup> Sánchez Ramos, et al. Traducción automática: conceptos clave, procesos de evaluación y técnicas de posesición. Comares, 2020.

<sup>50</sup> 冯志伟. (2010). 基于语料库的机器翻译系统. 术语标准化与信息技术, 计算机语言学. 第 1 期. (Feng Z. w. (2010). Sistema de traducción automática basado en corpus. Estandarización terminológica y tecnología de la información, Lingüística informática. No. 1)

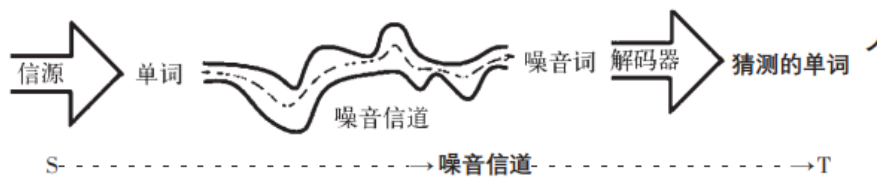


图1 噪音信道模型

Los sistemas de TA basados en estadística necesitan hacer pruebas y entrenamientos en cuanto a corpus. Según la figura 6, se ve que el sistema de traducción convierte el corpus bilingüe en un corpus estandarizado a través de transformarlo, preprocesa el corpus para extraer las palabras y hacer que las palabras bilingüísticas coincidan entre sí a fin de formar un corpus bilingüe con palabras alineadas. Para el corpus, el sistema realiza nuevamente el análisis sintáctico y la extracción de características de reordenación, así se puede obtener un corpus monolingüe, alineación de palabras y análisis sintáctico.

De acuerdo con Sánchez (2020), se resumen las ventajas y desventajas de los sistemas basados en estadísticas. Por un lado, en cuanto a las ventajas, el uso de corpus bilingüe en este tipo de sistemas garantiza la calidad de la traducción. En comparación con el sistema basado en reglas, se calcula por medio de fórmula y se analiza el corpus, así la traducción es más exacta y más rápida. Además, la capacidad de conocimiento y habilidad de aprender de este tipo de sistema es más poderosa. Por otro lado, debido a su teorema de funcionar, este sistema depende mucho en los recursos electrónicos. La condición para traducir bien consiste en que en este ámbito hay materiales, estudios, investigaciones, y corpus para analizar y calcular. Entonces, si no hay disponibles cursos para formular un corpus, es difícil llevar al cabo la traducción y la calidad de los textos traducidos será limitada.

Al mismo tiempo, el proyecto Moses, EuroMatrix, y MTradumática consisten en ejemplos de los sistemas de traducción automática basados en estadística.

### Figura 6

*Proceso de formación de los sistemas de TA basados en estadística (Feng Zhi Wei, 2010)*

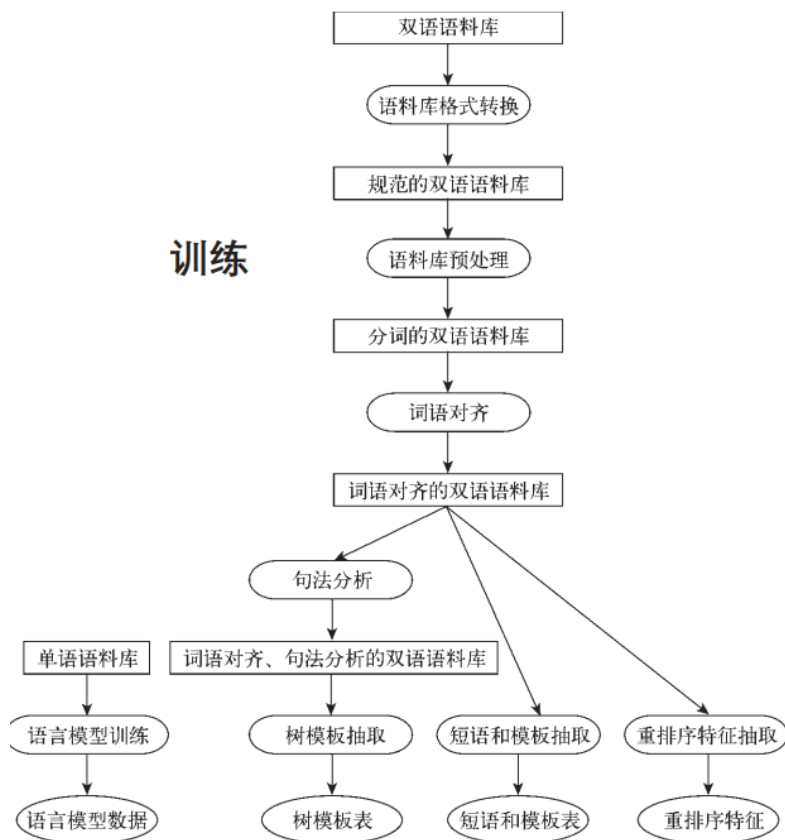


图3 统计机器翻译的“训练”流程

### Sistema de traducción automática basados en ejemplos

La idea de sistema de TA basado en ejemplos fue propuesta por primera vez por Nagao Maokoto (1984).<sup>51</sup> De acuerdo con Feng (2010), en el sistema basado en ejemplos, el conocimiento principalmente proviene de corpus de ejemplos de traducción bilingüe. El teorema de función consiste en que en el corpus se mantienen una oración del idioma de origen y la oración traducida. Cuando se introducen palabras u oraciones en este sistema, el sistema va a hacer un análisis con las en corpus y encontrar una traducción con más probabilidad. Desde el punto de vista de Sánchez (2020), los tres principales procesos de este tipo de sistema consisten en:

- a. La correspondencia de los segmentos en la lengua origen con los ejemplos existentes extraídos del corpus paralelo alineado.
- b. La alineación de los segmentos origen con sus correspondientes traducciones.
- c. La combinación de los mismo para generar la traducción del texto en la lengua meta. (Sánchez, 2020)<sup>52</sup>

Para mejorar la traducción de sistemas de TA basados en ejemplos, Feng (2010)

<sup>51</sup> Nagao, M. (1984). A framework of a mechanical translation between Japanese and English by analogy principle.

<sup>52</sup> Sánchez Ramos, et al. Traducción automática: conceptos clave, procesos de evaluación y técnicas de posesición. Comares, 2020.p34.

recomienda ejercer investigación y estudios en cómo realizar correctamente la alineación automática bilingüe, establecer un mecanismo de búsqueda efectivo de coincidencia de ejemplos, generar traducciones correspondientes a las oraciones del idioma de origen de acuerdo con el corpus de ejemplos.

En base a Feng (2010), actualmente existen los siguientes sistemas basados en ejemplos en el mundo: Sistema de traducción automática multimotor (Multi-engine Machine Translation) PAN-GLOSS de la Universidad Carnegie Mellon, EEUU; Sistemas MBT1 y MBT2 de Makoto Nagao y S. Sato, Universidad de Kioto, Japón, Sistemas E-TOC y EBMT de ATR, el Laboratorio de Japón de Investigación en Traducción de Lenguas Habladas y Comunicación; Sistema automática japonés-chino basado en ejemplos, Universidad de Tsinghua; Sistema “Daya”, el Instituto de Tecnología de Harbin y la Universidad de Tsinghua.

### **Sistemas de traducción automática híbridos**

De acuerdo con Sánchez (2020), los sistemas de traducción automática híbridos consisten en sistemas que combinan los sistemas basados en estadísticas y los sistemas basados en reglas. Desde el concepto de este sistema, es cierto que es un paso innovador para compensar las desventajas relativas del sistema basado en reglas y el basado en estadística. Como propone Li Yegang, Huang Heyan, Feng Chong y Su Chao (2015), hay diferentes sistemas de TA, pero cada uno tiene su propio inconveniente difíciles de superar. Al mismo tiempo, “cada uno se centra en una perspectiva y un aspecto diferentes del problema de la traducción”.<sup>53</sup> Entonces, es muy aconsejable dedicar a la investigación de combinar los métodos de sistemas para elevar la calidad y capacidad de traducción.

### **Sistemas de traducción automática basados en redes neuronales**

Los sistemas de TA basados en redes neuronales están encaminados a realizar la traducción del idioma de origen al idioma de destino por medio de redes neuronales. Este tipo de sistema es diferente en modelo de traducción con otros tipos de sistemas. Como explican Li Yachao, Xiong Deyi, Zhangmin (2018)<sup>54</sup>, este sistema utiliza el modelo llamado Continuos Space Representation para representar palabras, frases y oraciones. Se puede realizar la asignación del idioma de origen al idioma de destino por medio de la red neuronal sin necesidad de alineación de palabras y extracción de reglas. Existen dos modelos parecidos en el teorema, que uno es End-to End Model, propuesto por Google Translate, que otro es Encoder-Decoder Model, propuesto por la Universidad de Montreal.

---

<sup>53</sup> 李业刚,黄河燕,史树敏,冯冲 & 苏超.(2015).多策略机器翻译研究综述. 中文信息学报(02),1-9+23. (Li Yegang, Huang Heyan, Shi Shumin, Feng Chong y Su Chao. (2015). Una revisión de la investigación en traducción automática multi-estrategia. Revista china de informática (02), 1-9+23.)

<sup>54</sup> 李亚超,熊德意 & 张民.(2018).神经机器翻译综述. 计算机学报(12),2734-2755. (Li Yachao,Xiong Deyi y Zhang Min. (2018). Una revisión de la traducción automática neuronal. Revista de ciencias de la computación (12), 2734-2755.)

En este párrafo, se discute la diferencia entre el sistema de TA basado en estadística y el basado en las redes neuronales. El punto mismo de los dos sistemas consiste en que la idea central de los dos sistemas es tomar la traducción como resolución de problemas probabilísticos. En primer lugar, los modelos de la alineación de palabras son diferentes. En el sistema de TA basado en estadística, se aplican “la alineación de palabras y la correspondencia” (Sánchez, 2020). Pero, en el sistema basado en redes neuronales, no hace falta los pasos para realizar la alineación de palabras sino se puede realizarla por medio del modelo Attention Mechanism. En segundo lugar, los dos sistemas tienen diferentes ventajas en cuanto a la calidad de traducción. El sistema basado en redes neuronales puede gestionar los problemas de secuenciación compleja y traducción a larga distancia, mientras que el sistema basado en estadística puede mantenerse bien fiel al texto original.

De acuerdo con Li, Xiong, Zhang (2018), se resumen los desafíos que está enfrentando el sistema de TA basado en redes neuronales:

- 1) **Glosario y frase limitados:** En el proceso de generación del texto del idioma de origen, el sistema requiere un glosario de tamaño fijo bilingüe, lo que lleva dificultad en la traducción de traducir las frases muy largas.
- 2) **Falta de conocimientos avanzados:** En este sistema se aplica el entrenamiento de datos bilingüe, sino se requieren conocimientos avanzados como: corpus monolingüe a gran escala, corpus etiquetado, diccionarios bilingües, etc. Excepto, las características estructurales de este sistema brindan dificultades a la aplicación de recursos externos. Es cierto que los recursos y conocimientos avanzados puede mejorar en gran medida la calidad de las traducciones.
- 3) **Mecanismo de Atención:** El mecanismo de atención consiste en una característica destacable del sistema basado en redes neuronales. En el proceso de generación del idioma y destino, aparece el problema que ignora la información histórica de atención. Además, también existe el problema de la traducción excesiva y insuficiente (Over Translation y Under Translation).
- 4) **Arquitectura única:** Según lo explicado en el apartado anterior, la TA neuronal está basada en decodificadores-codificadores (Encoder-Decoder Model), así este modelo solo puede capturar información de palabras a nivel de oración. (Li, Xiong, Zhang, 2018)

A pesar de varios aspectos limitados, los sistemas de TA basado en redes neuronales disponen de muchas ventajas. En primer lugar, los textos traducidos por medio de estos sistemas disponen de alta calidad y exactitud. En segundo lugar, este tipo de sistema puede gestionar pares de lenguajes, de forma más flexible, más lógica. Por último, con la gran capacidad de motor y arquitectura, este tipo de sistemas viene con alta eficiencia, ahorrando mucho tiempo y coste. No solo se puede traducir palabras por palabras, sino es capaz de traducir las conjugaciones, declinaciones, etc.

Actualmente hay muchos software y plataformas en que se aplica el sistema de TA basado en redes neuronales. Se enumeran de acuerdo con Sánchez (2020): Google

Translate, DeepL, OpenNMT, Marian, Nematus, Sockeye, Nural Monkey, ect.

### **Clasificación de los sistemas**

En el apartado de los sistemas de traducción automática, están divididos en cinco diferentes sistemas de acuerdo con la tecnología aplicada. Aquí se introduce otros modelos para clasificar los sistemas de TA. De acuerdo con Pilar Hernández Mercedes (2002), desde el punto de vista idiomático, los sistemas de TA se pueden clasificar en “bilingüe y multibilingüe”(Mercedes, 2002). La diferencia entre los dos tipos radica en que el sistema puede traducir un idioma de origen a otra lengua de destino o puede traducir varias a varias. En cuanto a los sistemas de TA multibilingüe, se abarcan DeepL Translate, Google Translate, Tencent Translate, Youdao Translate, iPhone Translate, ect. Y el software llamado XiyuZhushou es un ejemplo de sistema bilingüe, que solo puede traducir entre español y chino. Como expone Mercedes (2002), de acuerdo con la dirección, se puede dividir en sistemas de TA unidireccional o bidireccional. Es fácil entender que se refiere a si el sistema puede soportar la traducción de un par de lenguas mutuamente.

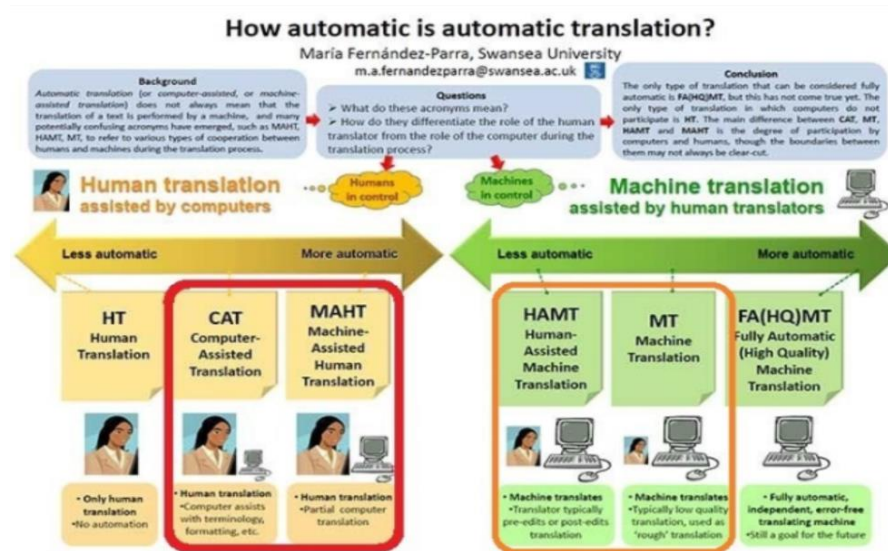
Además, los sistemas de traducción automática se pueden clasificar en sistemas de traducción de TA completamente automática, sistemas de TA asistida por humanos y sistemas de traducción humana asistida por ordenador (Mercedes, 2002). Es una clasificación importante en base al grado de automatización. En cuanto al primer tipo, el sistema realiza la traducción sin ninguna intervención humana. En cuanto al segundo tipo, el sistema realiza la traducción con la ayuda y asistencia humana. En cuanto al último, el sistema realiza la traducción con la ayuda de ordenador. Cabe destacar que el sistema de traducción humana asistida por ordenador está diseñado a ayudar a los traductores a través de proporcionar corpus, glosario, perfección, diccionario, información gramatical, etc.

En la figura 7, desde la derecha a la izquierda, están traducción humana, traducción asistida por ordenador, traducción humana asistida por máquina, traducción automática asistida por humanos, traducción automática y traducción automática completamente. El orden desde la derecha a la izquierda, la traducción es más automática, por lo contrario, desde la izquierda a la derecha, la traducción es menos automática. En esta figura se ve muy claramente el grado de automatización de diferentes traducciones. Entre los seis tipos de traducción, la HT, CAT, MAHT (se ven los nombres completos en la figura) están controladas de los humanos, mientras que, en la parte verde, la HAMT, MT, FAMT (se ven los nombres completos en la figura) están controladas de las máquinas. En la parte naranja, correspondiente a los tres tipos de traducción, la traducción humana es una traducción sin automatización; la traducción asistida por ordenador consiste en traducción con la ayuda de terminología, formato, etc. de ordenador; la traducción humana asistida por máquina es una parcial automática. En la parte verde, correspondiente a los tres tipos de traducción, la traducción automática asistida por humanos es una especialmente con preedición o postedición; la traducción automática suele ser de baja calidad y se toma

como traducción en bruto; la traducción completamente automática se considera un objetivo en el futuro, donde no hace falta la corrección y asistencia humana.

**Figura 7**

*How automatic is automatic translation? (María Fernández-Parra, sin fecha)<sup>55</sup>*



En la figura 8, según María Fernández-Parra (2009)<sup>56</sup>, la CAT trata de computer-asisted translation, en otras palabras, traducción humana asistida por computadora (TAC). Se presenta un proceso completo de TAC. De primero, antes de la traducción, cuando un traductor recibe el trabajo de traducción desde el cliente, puede analizar la cantidad de las palabras del texto y hacer una organización de tiempo. Después, puede extraer glosarios desde el texto e ingresa el texto en la computadora para hacer una traducción primaria. Con la memoria y glosario, el traductor puede corregir y ajustar la traducción. Luego de terminar la traducción, se puede usar la función de revisión de texto para examinar la calidad y terminologías.

Desde mi punto de vista, se trata de un tipo de traducción muy útil. Por un lado, la traducción primaria de la computadora puede ayudar al traductor eliminar las dificultades de la comprensión de estructura gramatical y sintáctica y montón de terminología. Aunque la computadora puede traducir algunos términos de manera incorrecta, el traductor puede comparar significado de contextos a fin de encontrarlos. En el mercado, la TAC es popular y se aplica mucho en diferentes ámbitos, como el estudio académico, la investigación y aplicación diaria. Actualmente existen las herramientas de traducción asistida: SDL Trados Studio, memoQ, Wordfast (Classic y pro), Déjà Vu, Across, Café Tran Espresso, OmegaT, Similis, Wordfast Anyway, MateCat, SmartCAT, Virtaal, Pootle, Lokaliza, GlobalSight, gtranslator, XTM Cloud, Star Transit, Poedit, Meta Taxis,

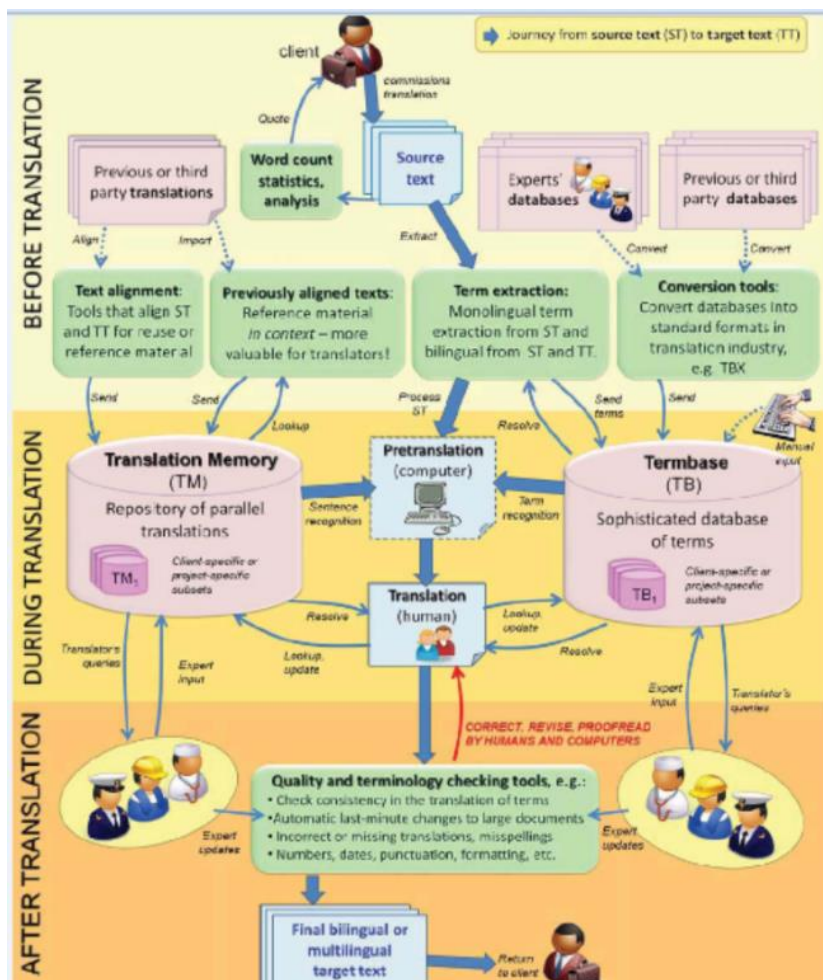
<sup>55</sup> Citado desde los cursos del profesorado Manuel Mata, no se encuentra la fecha de esta foto

<sup>56</sup> Fernández-Parra, M. (2009). The workflow of computer-assisted translation tools in specialised translation. In Reconceptualizing LSP: Online proceedings of the XVII European LSP Symposium.

Paraphrase, etc.

**Figura 8**

*A simplified but typical workflow with CAT tools (María Fernández-Parra, 2009)*



A continuación, en cuanto a la clasificación de sistemas de traducción, de acuerdo con Mercedes (2002)<sup>57</sup>, los sistemas se pueden dividir en “sistemas de traducción directa”, “sistemas de transferencia” y “sistemas interlingua”. Esta forma de clasificación está basada en el diseño y la estrategia general de funcionamiento. En primer lugar, en los sistemas de traducción directo, se traducen palabras por palabras, solo “reemplazando palabras de la lengua origen por palabras de la lengua destino” (Mercedes, 2002). Además, en este tipo de sistema se aplican las listas de palabras, diccionarios, y glosarios como los recursos electrónicos. Entonces, la unidad de esta traducción consiste en palabras, sin considerar el contexto, la gramática, la sintaxis, la conjugación, el tono y la sensación, etc. Así la calidad de traducción directa no es muy alta, y no se puede traducir de manera muy exacta y rápida. El idioma de origen y el idioma de destino son productos de diferentes culturas, por lo que la estructura y la forma de expresión son muy diferentes. De esta manera, es imposible llevar al cabo la traducción adoptada a la cultura.

<sup>57</sup> Mercedes, P. H. (2002). En torno a la traducción automática. *Cervantes*, 1(2), 101-117.



En segundo lugar, con respecto a los sistemas de transferencia, según Mercedes (2002), el proceso de traducción del texto está dividido en tres fases. De primero, se trata de análisis del idioma de origen “con el fin de obtener una representación intermedia con formación sintáctica y/o semántica.” (Mercedes, 2002) y con el motivo de recopilar las características de la lengua de origen para hacer un análisis. Así se puede decir que es una preparación antes de traducir. De segundo, la segunda fase consiste en la transferencia de las lenguas. El objetivo de transferencia es traducir los términos y ajustar las estructuras de la lengua de origen a través de aplicar las reglas. Por último, la tercera fase consiste en la generación. Se trata de la reconstrucción de la traducción en base a los datos obtenidos desde las dos fases arriba. Cabe mencionar que estos sistemas de transferencia se corresponden a sistemas de TA en base a reglas, sistemas de TA en base a ejemplos, sistemas de TA en base a estadística.

En tercer lugar, los sistemas interlingua está basado a una representación. Como enseña Zapata (1995), “La finalidad es que cualquier texto en cualquier lengua origen se represente con interlingua y que este lenguaje sea lo suficientemente potente como para representar cualquier tipo de información relevante para traducir”<sup>58</sup> Este teorema es muy parecido a la de sistemas de TA basados a redes neuronales.

En resumen, en este apartado se introducen diferentes maneras para clasificar los sistemas de traducción automática. Es importante para el traductor preparar la traducción y realizarla. Por un lado, a través de conocer los sistemas de TA, se puede conocer diferentes características de las herramientas en el mercado. El traductor puede elegir el instrumento adecuado según sus ventajas y desventajas. De acuerdo con la longitud de textos, tipologías, ámbito de terminologías, el traductor puede saber cómo realizar una traducción con más alta eficiencia. Por otro lado, es la teoría necesaria para los que se dedican a la traducción e interpretación, por lo que la tarea de traductores e intérpretes no se limita a los textos, sino tiene su papel en el desarrollo del estudio pertinente. Actúan como puente de comunicación de diferentes culturales y diferentes países, donde el nivel de desarrollo científico es diferente. La investigación y estudio de este tema tienen las oportunidades de ilustrar ingenieros, científicos, etc. Así, mantener una visión mundial y general no solo favorece la traducción, sino favorece el avance de tecnología. El tema de esta tesis se concentra en el traductor iFLYTEK, que consiste en una traducción automática con sistemas de TA basados en redes neuronales. La comprensión de la tipología y sus características puede contribuir a la evaluación de la traducción de los textos en el ámbito sanitario.

## **2.2 La traducción automática en el ámbito sanitario**

En el apartado anterior, se explican la definición de la traducción automática, su historia de desarrollo y la tipología de TA. En esta parte se propone la aplicación de la TA en el ámbito sanitario.

---

<sup>58</sup> I. ZAPATA, “Un recorrido por algunos principios de la Traducción Automática” en Enseñanza de lenguas y traducción con ordenadores, ed. G. Ruipérez, Madrid: Ediciones pedagógicas, 1995, 87-126.

Como descubre Lorena Turjillos-Yébenes y Ana Muñoz-Miquel (2022), no existe mucho estudio dedicado al análisis de TA en este ámbito muy específico, pero, de verdad hay varios estudios apuntados a la traducción automática estadística.

En las últimas décadas, en China cada vez más estudiosos empiezan a dedicarse a la investigación de TA en el ámbito sanitario:

- 1)张涛, (2011), 面向特定领域的统计机器翻译研究与应用 (Zhang Tao, Investigación y aplicación de la traducción automática estadística para un dominio específico)
- 2)刘洪涛, (2020), 基于生物医学语料的神经机器翻译系统研究 (Liu, Hongtao, Investigación sobre un sistema neuronal de traducción automática basado en un corpus biomédico)
- 3)张邵源, (2021), 面向生物医学领域的神经机器翻译领域适用方法研究 (Zhang, Shaoyuan, Investigación sobre métodos aplicados en el campo de traducción automática neuronal en el ámbito de la biomedicina)
- 4)陈相 (2008), 面向生物医学领域的双语对齐技术研究 (Chen Xiang, Investigación sobre tecnología de alineación bilingüe para el ámbito biomédico)
- 5)郭子凯, 陈晨, & 刘曼玲. (2022). 译者主体性视角下医学学术文本“机器翻译+译后编辑”中的策略研究 (Guo Zikai, Chen Chen, & Liu Manling, Investigación de estrategias en "traducción automática + postedición" de textos académicos médicos desde la perspectiva de la subjetividad del traductor)
- 6)盛诗梦, 丁皓, 徐欣茹, 丁思吉, & 夏冬阳. (2022). 注意力机制在医学上的应用综述 (Sheng Shimeng, Ding Hao, Xu Xinru, Ding Siji, & Xia Dongyang. (2022). Una revisión de los mecanismos atencionales para aplicaciones médicas.)
- 7)曾江霞. (2020). 多模态大数据语境下科技术语翻译标准分析——以新冠肺炎和新冠病毒术语翻译为例 (Zeng Jiangxia. (2020). Análisis de estándares de traducción de términos científicos y técnicos en el contexto de big data multimodal: ejemplos de traducción de los términos neotripanosomia y virus Xinkang. Términos científicos y técnicos chinos)

Después de analizar estas investigaciones, se encuentra que el enfoque de estos autores no está en el desarrollo o investigación de la teoría de la traducción automática, sino de acuerdo con los sistemas originales, sistemas basados en reglas, sistemas basados en estadística, sistemas basados en ejemplos, sistemas basados en redes neuronales, están en la innovación y diseño para mejorar la calidad de TA.

Desde los estudios enumerados arriba, podemos saber que, aunque la TA en el ámbito sanitario no se convierte madura, muchos autores insisten realizando estudios para hacer innovación y seguir adelante, lo que puede contribuir al desarrollo de la TA. La capacidad de TA en campos específicos, no limitados en el ámbito sanitario, al fin al cabo, depende del nivel de tecnología y el diseño de ingenieros y científico. En esta época, en la globalización, los intercambios económicos, políticos, médicos, sanitario, etc. están convirtiéndose cada más frecuentes. Así surge una gran necesidad de TA en los ámbitos

debido a la dificultad y barreras de culturas y lenguas.

En China no existen muchos estudios e investigaciones dedicadas a la TA en el ámbito sanitario. No obstante, varios autores hacen propuesta al diseño de la TA en campos especializados. Desde mi punto de vista, estos varios establecen base teórica para la realización y la aplicación en la realidad. A lo siguiente, se presentan varios sistemas diseñados para realizar la TA.

De primeo, de acuerdo con Zhang Tao (2011)<sup>59</sup>, él diseñó un sistema de TA basado en estadística a fin de facilitar la lectura y traducción en ámbito médico. En su sistema, se aplican las siguientes herramientas: herramienta de alineación de correspondencia de palabras: giza-pp-v1.0.5.tar.gz; herramienta de segmentación de palabras chinas: ICTCLAS; herramienta de capacitación de modelos de idiomas: SRILM; analizador de sintaxis en inglés: Stanford parser; herramienta de segmentación y preprocesamiento de palabras en inglés: tokenizaE.perl.tmpl; Analizador de sintaxis de dependencia china: ctbparser. Todas estas herramientas pueden satisfacer la necesidad de generación del idioma de origen y del idioma de destino. Esta investigación se dedica al par de lenguas: chino-inglés. Pero se puede lograr mismo efecto reemplazando estas herramientas de inglés por las de español, así es posible diseñar un sistema espacial para español-chino. Aquí se pone un recurso muy útil para consultar los materiales y conocimientos en el ámbito sanitario. Se trata de Word Core Courses of Chinese Medicine especially, Federation of Chinese Medicine Societies (2012).

Según Zhang (2011), en este sistema diseñado, se abarcan los siguientes recursos en el ámbito sanitario: 2626 corpus paralelos, 1900 plantillas de oraciones y 12000 plantillas de frases y palabras. Además, se utiliza el sistema de SRILM para el entrenamiento del modelo del lenguaje.

### **Figura 9**

*Plantilla del sistema de traducción (Zhang Tao, 2011, p.44)*

---

<sup>59</sup> 张涛, (2011), 面向特定领域的统计机器翻译研究与应用, 昆明理工大学 (Zhang, Tao, (2011), Investigación y aplicación de la traducción automática estadística para un dominio específico, Universidad Tecnológica de Kunming.)

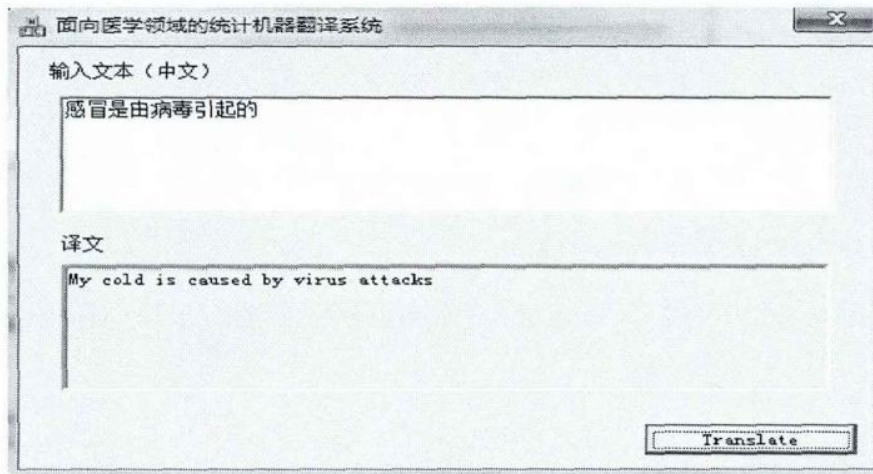


图 5.2 系统翻译界面

En la figura 8, se ve muy claramente la plantilla del sistema diseñado de Zhang Tao(2011) para la TA en el ámbito sanitario. La página de este sistema es muy simple y sencillo. En el cuadro arriba se puede introducir lo que quiere buscar en chino sobre la ciencia médica, como el medicamento, síntomas, términos o frases para consultar al médico, y después, el motor va a realizar la alineación, correspondencia y generación, al fin, en el cuadro debajo, se presenta la frase u oración traducida.

Sin duda, se trata de una innovación. Aunque según los recursos aplicados, la base de datos, terminologías, y corpus no abarca mucha cantidad de recursos electrónicos, el teorema y el punto de salida tiene base teórica y funciona este sistema. Pero, este sistema no se pone en la realidad, que es un diseño de graduación. Por lo tanto, este sistema está en la fase primaria, y se perfeccionará por medio de aplicación en las prácticas. Como no se puede traducir con este sistema, no se tiene oportunidad de hacer investigación y estudio para evaluar la calidad y analizar errores de este sistema. Al fin y al cabo, el diseño del sistema de TA en el ámbito sanitario significa que el público ya presta cada vez más atención a la TA en los textos especializados.

De segundo, Liu Hongtao (2020) lleva al cabo una investigación sobre un sistema neuronal de traducción automática basado en corpus biomédico. Este estudio está basado en la fundación de corpus en el ámbito sanitario y establecimiento de sistema. Pero, este sistema está apuntado al par de idiomas chino-español. En cuanto a la adquisición de corpus médico, se aplica un sistema para capturar todos los artículos médicos más autorizados y orientados en los proyectos de la Red de Revistas Médicas Chinas y la Fundación Científica, y por fin extraje todos los documentos relevantes de 2005 a 2018.

Aquí hace falta introducir la Red de Revistas Médica Chinas (China Medical Journals Network). Se trata de una plataforma donde se reúnen investigaciones y estudios avanzados en el ámbito médico, convirtiéndose en la más grande red de revistas profesionales médicas. Su sitio web oficial: <https://www.medjournals.cn/index.do>

De acuerdo con Liu (2020), su investigación agrega tecnología de desambiguación semántica y diccionario profesional biomédico a la arquitectura del transformador con la Self-Attention como núcleo. Además, realiza entrenamiento de traducción biomédica basado en el corpus paralelo bilingüe chino e inglés capturado. Después de todo, él diseña un sistema de traducción neuronal biomédico llamado BioNMT System. El resultado y efecto de este diseño de sistema es ideal en comparación con traductor tradicional. La calidad de traducción de textos del idioma es mejor con la exactitud de correspondencia de términos y la eliminación de ambigüedad.

Aquí se pone un ejemplo muy destacable de la traducción automática en el ámbito sanitario. La plataforma de ATAN (sitio web: <https://fanyi.atman360.com/>) se ha lanzado en el público y ahora es popular por las empresas. Tiene tres funciones principales: traducción de palabras, traducción de documentos, y traducción CIOMS. En cuanto a la función de traducción de palabras, esta plataforma es diferente con otras plataformas, como Google Translate, DeepL Translate, centrandose en la traducción precisa en siete subdivisiones: genéricos, farmacovigilancia, dispositivos médicos, CMC, clínica, no clínica y reglamentaria sino en campos generales en el ámbito sanitario. Además, esta plataforma soporta la memoria de corpus almacenada en el sistema para que el motor pueda aprender, haciendo el resultado más adoptado a los hábitos lingüístico. Con respecto a la traducción de archivos, esta herramienta admite la conversión y traducción de documentos en 9 formatos y la carga por lotes y el procesamiento de archivos, elevando la eficiencia de rendimiento. Al mismo tiempo, se puede realizar la lectura cruzada y revisión en línea del texto original y el texto traducido, lo que hace posible mejorar la calidad ahorrando el tiempo y la energía. Este motor se adapta a diferentes escenarios en la industria farmacéutica como farmacovigilancia, solicitud de registro, consultas médicas, investigación clínica, marketing, revisión de literatura, compañías de traducción farmacéutica, integración sistemas médicos y otros escenarios. Es por todo eso que puede mejorar en gran medida la eficiencia de la conversión de idiomas y reducir efectivamente los costos de mano de obra. No obstante, esta plataforma hasta ahora solo está abierta a las empresas, entidades o compañías, sino se puede usar por individuales con el objetivo de hacer investigación o realizar la traducción. Pero, de acuerdo con sus usuarios objetivos, la calidad de traducción sería de alto nivel. También se trata de un área muy interesante para profundizar.

### **2.3 Evaluación de la TA**

La evaluación de la traducción automática consiste en un tema muy importante desde el punto de vista académica y científica. La evaluación precisa de los textos traducidos es la base para la investigación y desarrollo de los sistemas de TA.

#### **Figura 10**

*Diagrama del ciclo de desarrollo de la TA (Qin Ying, 2015)*

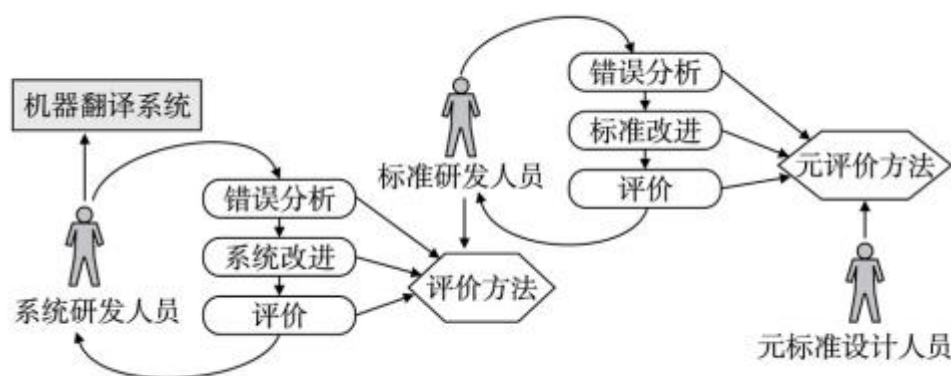


图1 机器翻译开发周期图

Desde la figura 10, se ve que el desarrollador de las métricas de evaluación es responsable analizar el error de la traducción realizada por TA y mejora las métricas. Los desarrolladores de sistemas de TA realizan análisis de errores producidos y mejora el sistema en base a las causas. A partir de esto, se puede saber que las métricas efectivas de evaluación de TA son un catalizador para que la tecnología y sistema avancen adelante.

Como dice Sánchez (2020), “La evaluación propiamente dicha depende de diversos factores y puede realizarse con diferentes propósitos”<sup>60</sup> Es un estudio subjetivo en algún sentido, es que diferentes grupos tienen diversas preferencias a la traducción. Existen diferentes estándares y métricas de diferentes autores en cuanto a la evaluación. Generalmente, según Sun Yiqun y Zhou Minkang (2017), la evaluación de la calidad de una traducción debe tener en cuenta tres factores principales: de primero, vocabulario, es decir, debe prestar atención a significado de palabras, retórica, términos técnicos y uso del dialecto; de segundo, gramática, en otras palabras, debe prestar atención al uso gramáticos; de tercero, discursos, es decir, debe prestar atención a la cohesión, coherencia, intención, aceptabilidad, información, contexto e intertextualidad de la traducción.<sup>61</sup>

Generalmente, existen dos diferentes métodos para hacer la evaluación de TA: la evaluación automática y la evaluación manual. El primer tipo se realiza por máquinas sin intervención humana, como las métricas automáticas como CDER (Cover Disjoint Error Rate), BLEU (Bilingual Evaluation Understudy), TER (Translation Edit Rate), WER (Word Error Rate), GMT (General Text Matcher), METEOR (Metric for Evaluation of Translation with Explicit Ordering) y semiautomáticas (HTER, Human-mediated Translation Error Rate). El segundo tipo se realiza por humanos en caso de disposición de textos paralelos de referencia de par de lenguas, o en caso de ser traductores profesionales o bilingües.

<sup>60</sup> Sánchez Ramos, et al. Traducción automática: conceptos clave, procesos de evaluación y técnicas de posesición. Comares, 2020.

<sup>61</sup> 孙逸群, & 周敏康. (2017). 机器翻译质量综合评价方法研究. 中国科技翻译, 30(2), 20-24.(Sun, Yijun, y Zhou, M.K.. (2017). Investigación sobre el método de evaluación integral de la calidad de la traducción automática. Chinese Scientific and Technical Translation, 30(2), 20-24.)

### 2.3.1 Evaluación automática de la calidad de TA

En este apartado se propone la breve historia del desarrollo, las ventajas y desventajas, y características de la evaluación automática de la calidad de TA. Además, se aplica las métricas de BLEU, METOR y TER para poner en la práctica la evaluación de dos textos paralelos bilingüe a fin de resumir el uso y sus rasgos. De acuerdo con Silvia Esteban Hernández (2020), “Las métricas de evaluación automáticas de TA consisten en una serie de algoritmos y procesos que calculan puntuaciones numéricas basándose en ciertos aspectos de la calidad, que pueden variar según la métrica.”<sup>62</sup>

En primer lugar, de acuerdo con Sánchez (2020), se resume la historia de evaluación para tener un concepto de la identificación:

- 1) En 2002: Pepinei, Roukos, Ward y Zhu diseñaron la métrica llamada BLEU (Bilingual Evaluation Understudy).
- 2) En 2003: el National Institute of Standards and Technology lanzó un programa de evaluación de TA, el DAPRA TIDE basado en la métrica BLEU, en EE.UU.
- 3) En 2004: Alon Lavie desarrolló la métrica METEOR (Metric for Evaluation of Translation with Explicit Ordering) en el Language Technologies Institute, Pittsburgh, EE.UU.
- 4) En 2006: Mathe Snover desarrolló la métrica llamada TER (Translation Edit Rate), en la Universidad de Maryland, en EE.UU.
- 5) En 2006: el National Institute of Standards and Technology diseñó la métrica NIST MT.
- 6) En 2007: surgen diferentes iniciativas sobre las métricas automáticas.
- 7) En 2008: los primeros estudios sobre la comparación de las métricas automáticas se realizan por el National Institute of Standards and Technology, en EE.UU.<sup>63</sup>

Desde el proceso de desarrollo de las métricas de evaluación de TA, se encuentra que la fecha de inicio se centra en la década a principios del Siglo XX. Además, las métricas BLEU, METEOR, y TER se diseñan en EE. UU, lo que refleja el poder de tecnología del país. En China existen muchas herramientas de traducción automática, pero todavía no se desarrolla un programa para evaluar, así la evaluación de la traducción todavía depende de los humanos. Con esta investigación, se espera brindar ilustración y un poquito base teórico para mi país.

A lo siguiente, se analiza las ventajas y desventajas de la evaluación automática desde el punto de vista práctico, metodológico y científico. Las métricas pueden evaluar de manera rápida y de alta eficiencia, pero el nivel de exactitud está pendiente de discutir. En una investigación muy seria e importante, es mejor combinar la evaluación humana y evaluación automática.

---

<sup>62</sup> Esteban Hernández, S. (2020). Evaluación de la calidad de la traducción de motores de traducción automática neuronal en textos del ámbito jurídico.

<sup>63</sup> Sánchez Ramos, et al. Traducción automática: conceptos clave, procesos de evaluación y técnicas de posesición. Comares, 2020.p 44.

**Tabla 7***Ventajas y desventajas de la evaluación automática de la TA*

Ventajas	Desventajas
El rendimiento de las métricas es alto, ahorrando el tiempo y la energía.	Está limitada a evaluar los textos con referencia o textos bilingüe.
Es fácil de operar la evaluación automática, reduciendo el coste para los traductores profesionales.	Existe posibilidad que los textos de referencia no son de muy alta calidad.
Se presenta resultados de manera de datos y formatos, haciendo más visible la calidad de traducción.	Solo se puede evaluar según el nivel de similitud con texto de referencia sin reconocer el punto sobresaliente de traducción.
Se realiza la evaluación por medio de motor, haciendo más objetivo el resultado.	El sistema de métrica es rígido. En cuanto a una oración hay diferentes maneras para expresar, pero el sistema no puede reconocer, sino evaluar si es mismo o no con texto de referencia.
Existen diferentes métricas, y se pueden comparar los resultados de una misma traducción para obtener un resultado más preciso.	Los resultados tienen gran diferencia de acuerdo con la tipología de textos y campos. No se puede pensar y elegir. En la traducción se aplican diferentes técnicas, lo que contribuyen diferentes expresiones de una misma frase.

A lo siguiente, de acuerdo con Sánchez (2020) (citado desde Lavie, 2011), los conceptos característicos consisten en precisión, exhaustividad, y distancia de edición de levenshtein. Un ejemplo de las métricas basadas en la precisión es BLEU, un ejemplo de las basadas en la exhaustividad METEOR, y un ejemplo de métricas basadas en la distancia de levenshtein TER.

**Figura 11***Tipología de evaluación de traducción automática (Qin Ying, 2015)*



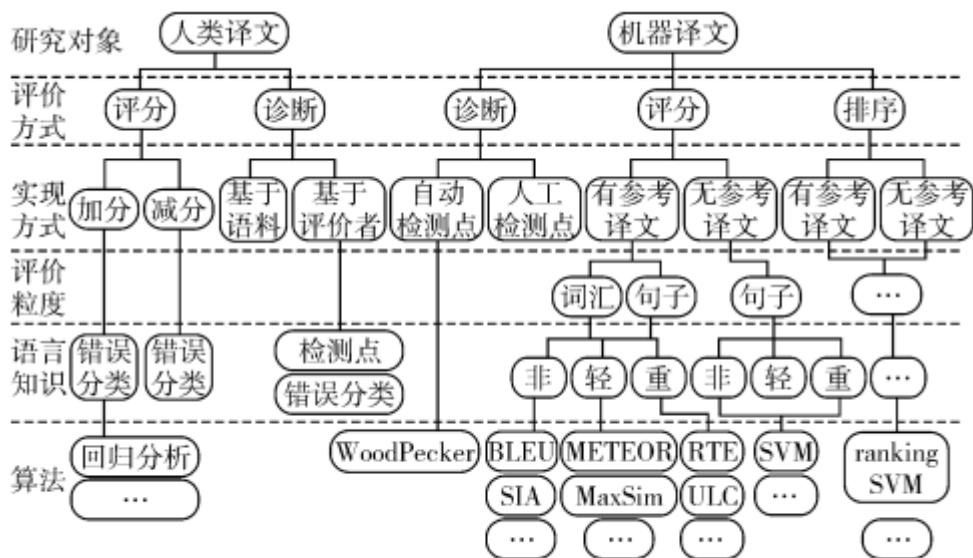


图2 翻译自动评价研究分类图谱

Desde la figura 11, se ven de manera clara el proceso de trabajo y la tipología de evaluación de traducción automática. Por un lado, en cuanto a la evaluación de traducción automática, los métodos para evaluar consisten en puntuación y diagnóstico. En el proceso de dar puntos, la implementación es sumar puntos y restar puntos. Y el traductor que realiza la evaluación es responsable de analizar los errores y clasificarlos, En el proceso de diagnóstico, se trata de un trabajo basado en corpus y revisores, y, además, hace falta clasificar los errores y hacer un resumen y análisis. En cuanto a la evaluación automática de la TA, los métodos para evaluar consisten en diagnóstico, puntuación y clasificación. En el caso de puntuación con traducción de referencia, la granularidad de la evaluación es vocabulario y oraciones, y en este proceso, BLEU, SIA, METEOR, MaxSim, RTE, ULC, etc., estas métricas realizan la evaluación de acuerdo con este método. Por otro lado, en el caso de puntuación sin traducción de referencia, la granularidad de la evaluación solo es la oración. La métrica SVM es un símbolo de este método. Al mismo tiempo, en la evaluación de TA, la métrica ranking SVM es un ejemplo del método de clasificación, sin importar tener traducción de referencia o no. Esta figura es muy sencilla, pero contiene casi la lógica básica de las dos diferentes maneras para evaluar la traducción automática.

Como el tema de esta tesis consiste en el análisis de la traducción de iFLYTEK en el ámbito sanitario, se aplicará una métrica para hacer la evaluación de la traducción automática. Entonces, se aplicará la métrica BLEU y aquí se analizan las características de BLEU.

Antes que nada, BLEU (Bilingual Evaluation Understudy) está basada en ngramas entre la traducción automática y la traducción de referencia, es decir la similitud entre sí. La precisión de similitud se puede calcular por medio de dividir ngramas comunes por ngramas candidata. Su sitio web es: <https://www.letsmt.eu/Bleu.aspx>. Hace falta seleccionar el archivo de textos originales, archivo de traducción humana de referencia, y archivo de traducción automática.

Pero esta métrica solo admite los documentos en formato de texto. Si los documentos son de pdt, word o ppt, es necesario cambiar el formato antes. Cabe mencionar que esta función es muy inconveniente, y es aconsejable que se desarrolle la función para admitir todos los formatos.

## Figura 12

*Ejemplo de Resultado de evaluación de traducción automática (BLEU, 2023)*

BLEU:	<b>0.34</b>			
Precision x brevity:	0.34 x 100.00			
Type	1-gram	2-gram	3-gram	4-gram
Individual	6.53	0.19	0.13	0.09
Cumulative	6.53	1.12	0.54	0.34
Export data	<a href="#">CSV</a>			

De acuerdo con la figura 12, se puede ver el porcentaje de la similitud de la traducción automática en comparación con la traducción de referencia. Además, también se ve cada oración en español, así como la traducción de referencia y la TA correspondientes. Con esta métrica, se puede hacer análisis de todas oraciones y obtener un resultado exacto de similitud. Desde la figura 11, se ve que el resultado es de 0.34. La similitud que presenta esta métrica está entre 0 y 1. Cuánto más cerca de 1, significa que es más similar la traducción automática con la traducción de referencia. Por lo tanto, en cuanto al resultado, se puede entender que la similitud de TA con la traducción de referencia es del 34%.

De acuerdo con Sánchez (2020), la métrica BLEU también cuenta con varias limitaciones. Por un lado, este sistema está diseñado para gestionar y generar corpus, así en caso de evaluación de oraciones individuales, la cantidad de análisis no es suficiente, contribuyendo la infidelidad de los resultados de la evaluación. Por otro lado, como el funcionamiento de BLEU depende de ngramas de traducción de referencia, no es capaz de reconocer los sinónimos, o expresiones diferentes con mismo significado. Sin embargo, las limitaciones no solo son de BLEU, sino de muchos sistemas automáticos de traducción o de evaluación, por lo que el uso de las herramientas debe salir de la necesidad y objetivo.

Según Angeru Mayor, Iñaki Alegria, Arantza Día de Ilarraza, Gorika Labaka, Mikel Lersundi y Kepa Sarasola (2009), por un lado, en los casos de seguimiento de los cambios incrementales de un sistema, comparación de sistemas que usen estrategias similares, optimización de los valores de los parámetros de sistemas estadísticos<sup>64</sup>, la métrica BLEU puede desempeñar su papel de manera muy adecuada. En cambio, si se

<sup>64</sup> Mayor, A., Alegria, I., de Ilarraza, A. D., Labaka, G., Lersundi, M., & Sarasola, K. (2009). Evaluación de un sistema de traducción automática basado en reglas o por qué BLEU sólo sirve para lo que sirve. *Procesamiento del Lenguaje Natural*, (43), 197-205.

realizan comparación de sistemas que usen estrategias diferentes, comparación de sistemas cuando el par de lenguas, el número de referencias o el tamaño de n-gramas es diferentes, identificación de mejoras de aspectos de la traducción que la métrica no modela bien o monitorización de mejoras que aparecen poco en el corpus de test, la métrica BLEU no puede desempeñar su rol muy bien. Así, de acuerdo con esos casos, es aconsejable preferir maneras adecuadas para hacer la evaluación y la investigación.

### **2.3.2 Evaluación humana de la calidad de la TA**

En este apartado se introducen las principales funciones, características, procesos y unos modelos de la evaluación humana. La evaluación humana se refiere a que es por los traductores profesionales o traductores bilingüe que llevan al cabo la evaluación de la traducción automática sin intervención y asistencia por herramientas u ordenadores.

De acuerdo con M. Isabel Diéguez M. Ileana Cabrera P. (1997)<sup>65</sup>, se presentan generalmente los procesos para hacer la evaluación. La evaluación consiste en una lectura contrastiva de los textos de fuente y de meta para si se traduce de manera completa todo el contenido del texto fuente en el texto meta. Esta fase también está encaminada a obtener información para resumir los errores, ambigüedad o inconsistencias.

La evaluación humana, también conocida como evaluación maunal (Sánchez, 2020) dispone de característica de objetividad. Es que se realiza por humanos, y cada persona tiene su preferencia en el arte o en el uso de términos o expresiones. Aunque hay modelo o principios para consultar, los comentarios y resultados de una misma traducción automática serán diferentes en algún sentido a través de evaluar por humanos. Sin embargo, esta manera tiene su propia ventaja. Por un lado, la evaluación automática con métricas se realiza en base de comparación con las palabras comunes, pero existe un problema, es difícil para evaluación automática reconocer la coherencia, la gramática, la sintaxis de una frase e incluso un párrafo. Es que la traducción no puede ser solo unas palabras sino un texto para expresar ideas así es coherente y correspondiente por todo el texto. En particular, en cuanto a español, en la parte gramática, los verbos tienen muchas conjugaciones según el tiempo y el tono. Después de estudiar diferentes métricas, los sistemas no están divididos según diferentes idiomas, lo que significa que la evaluación automática es como un formulario, solo haciendo análisis desde el punto de vista de matemática sino desde el punto de vista lingüístico. No presta atención a las características de cada idioma.

Al mismo tiempo, cuesta mucho para las métricas evaluarlo si es correcto o no. La evaluación humana es capaz de comprender y comentar la traducción de manera más flexible. Para expresar un significado hay diferentes métodos para expresar, en voz pasivo, impersonal, etc. Los traductores profesionales pueden hacer un comentario desde el punto de vista más amplio en base a todo el documento. Por otro lado, en algún sentido, la

---

<sup>65</sup> Diéguez M., M. I. , & Cabrera P., I. . (1997). Traducción humana o traducción automática: variables para una toma de decisiones eficiente. *Onomázein*, (2), 227 - 242.

evaluación humana tiene más espacio para mejorar la traducción. Según lo propuesto en el apartado arriba, para realizar la evaluación hace falta un texto original, un texto de referencia y un texto de traducción. Pero el objetivo de hacer la evaluación de un texto consiste en verificar la calidad suya y después considerar si aplicar la traducción en el estudio o en investigación. No todos los documentos tienen textos de referencia.

Entonces, si ya hay una versión muy perfecta de traducción, no son necesario traducir el documento otra vez por máquinas o herramientas y evaluarla con métricas. Por lo que, en la vida de estudio o diaria, excepto para hacer una investigación, la evaluación automática no es muy práctico. Aunque la evaluación humana es subjetiva en algún sentido, en base a modelos, puede realizar la evaluación de manera eficaz. Por otro lado, como se aplica un texto de referencia en la evaluación automática, se considera sin decir, que la calidad de la TA depende de la similitud con el texto de referencia. Sin embargo, el texto de referencia no es siempre perfecto, es que el texto de referencia se traduce por humanas u ordenadores asistidos por humanos. Así se puede decir que la calidad del texto de referencia a algún nivel es relativamente perfecta. Si una traducción automática es mucho mejor que el texto de referencia es difícil presentar un resultado correspondiente por la evaluación automática. Sin embargo, en este caso, la evaluación humana puede compensar este punto y dar un comentario adecuado, sin ser limitado a una referencia. No obstante, la evaluación humana también tiene desventajas. En primer segundo, como se realiza por traductores profesionales, el coste de mano de obra será alto. Es que el trabajo de hacer la evaluación no se puede llevar al cabo en un plazo corte. Los traductores necesitan revisar palabras por palabras, frases por frases. En segundo lugar, por misma razón, la energía y el rendimiento de humanos es limitada. Así es una gran dificultad para traductores evaluar un corpus de montón de palabras. Con un texto de palabras limitadas, las ventajas de la evaluación humana son muy destacable. Si es un estudio o un corpus grande, casi es imposible evaluar solo por medio de traductores profesionales. Por último, la evaluación de las métricas es más objetivo. Aunque en algún caso aparece resultado inexacto, se hace análisis cada frase y con figura y datos se hace más visible y claro el comentario de traducción.

Después de analizar las características de la evaluación manual, en este párrafo se propone los rasgos principales de la investigación de la evaluación de humanos. Según Wu Guangjun (2006), se resumen los siguientes puntos:

- 1) La mayoría de los modelos cuantitativos están basados en el micronivel del discurso;
- 2) La clasificación de los “errores” es dicotómica y compleja;
- 3) Muchos de los principios de referencia tienen en cuenta tanto las características del micro nivel como las del macro nivel del discurso, como la articulación y la coherencia;
- 4) Ninguno de los modelos define de manera clara el concepto general de evaluación de la calidad ni la tolerancia al error;
- 5) El concepto de error no está claramente definido, así no es fácilmente operacionar;
- 6) La mayoría de las teorías evitan el uso de los “criterios” sino aplican en su lugar “principios o directrices”.<sup>66</sup>

---

<sup>66</sup> 武光军. 2006. 《翻译质量评估: 论辩理论模式》评介[J]. 外语研究 (Wu, Guangjun. 2006. 2006. Revisión de la evaluaci

Hasta ahora, ya existen diferentes modelos para la evaluación humana. Y se analiza cada modelo y comparan las características de ello para saber cuándo elegirán qué modelo para hacer la evaluación.

### **Modelo “Xin”, “Da”, “Ya”**

De primero, se introduce el modelo de Fan Shouyi (1986). En el primer capítulo, se menciona que “信” (Xin), “达” (Da), “雅” (Ya) son tres principios básicos en la traducción. “Xin” se refiere a la presión. “Da” se refiere a la fluidez. “Ya” se refiere a la manera adecuada. Fan (1986) propuso por primera vez representar tres principios en una relación matemática. Se utiliza “F” para simbolizar “Xin”, “R” para “Da”, “S” para “Ya” del texto original y "S'" para “Ya” del texto traducido, se puede utilizar la siguiente fórmula para expresar la relación:  $I = (R / F) \cdot (S' / S)$  <sup>67</sup>

Como explica Zhang (2007), "F" y "S" pueden considerarse constantes "R" está más cerca de "F". Cuanto más cerca esté "S'" de "S", más cerca estará "I" (un exponente de lo cerca que está la traducción del original) de 1; no es posible igualar 1 exactamente. no es posible. El intervalo de S a 0 indica la posición exacta de este índice de género. <sup>68</sup> Se puede deducir que el resultado de los documentos argumentativos como ensayo, análisis será cerca al 1.

En cuanto a esta métrica, aunque existe esta fórmula, la lógica no es muy madura. Es que todavía no existen estándares y principios para evaluar “Xin”, “Da”, “Ya”. Así los tres parámetros todavía no se fijan y son subjetivas, es difícil garantizar la calidad de la evaluación en base a este modelo. Pero esta teoría se propuso en el siglo XX, a pesar de falta de base madura y suficiente, es resultado muy útil. Es que se trata de un paso innovador utilizando teorías de las matemáticas difusas para transformar la evaluación cualitativa de las traducciones expresada en una evaluación expresadas en “números”. (Zhang, 2007)

### **Modelo SAEJ2450**

De segundo, se introduce la aplicación de la métrica SAEJ2450 desarrollado por Society of Automotive Engineers (SAE International). Según Sánchez (2020), esta métrica se puede utilizar para hacer la evaluación tanto de la traducción humana como la traducción automática.

Como en la figura 11, ya se encuentra el método de evaluadores consiste en puntuación y diagnósticos de errores. De acuerdo con Sánchez (2020), esta métrica SAEJ2450 dispone

---

ón de la calidad de la traducción: un modelo teórico de argumentación [J]. Estudios de Lenguas Extranjeras)

<sup>67</sup> 范守义. 1987. 模糊数学与译文评价. 中国翻译 (Fan Shouyi. 1987. Matemáticas difusas y evaluación de la traducción. Traducción china)

<sup>68</sup> 张霄军. (2007). 翻译质量量化评价研究综述. 外语研究, (4), 80-84. (Zhang Xiaojun. (2007). Revisión de la investigación sobre la evaluación cuantitativa de la calidad de la traducción. Estudios de Lenguas Extranjeras, (4), 80-84.)

de las siguientes partes: de primero, los traductores pueden evaluar la exactitud en los términos, puntuación, orden de palabras y concordancia, omisión, sintaxis, ortografía. En cuanto a este punto, hay otros aspectos para prestar atención, como gramática, coherencia, expresión de arte, sinónimos, etc. Sánchez (2020) propone que los errores están clasificados en dos niveles, un nivel grave, y un nivel poco importante. Aquí se propone una duda. ¿Cómo juzgar si un error es grave o poco importante? Parece que es un poco subjetivo. Además, desde mi punto de vista, se puede mejorar en la generalización. Es que el problema gramático también es un error grave. En la lista de errores no se abarcan todos los errores frecuentes en la traducción, así el resultado de la evaluación de humanos puede ser influenciado en algún sentido.

Según la figura 12 y la figura 13, se puede hacer puntuación en cuanto a los errores aparecidos en la traducción. Pero aquí se propone una duda. En la figura 14, se introduce que la penalización por error grave en el aspecto de terminología consiste en 5, pero en la figura 15, se escribe que la penalización es de 4. No estoy segura si es que no entiendo claramente o es una ignorancia pequeña.

En resumen, esta métrica consta de muchos aspectos de errores y establecer un sistema para hacer puntuación. En algún sentido, es beneficioso hacer la evaluación manual más objetivo. Aunque en la lista falta algunos errores, y no hay criterios para juzgar el nivel grave de errores, esta métrica funciona de manera muy clara en el estudio y en la práctica. Por un lado, las dos tablas para hacer puntuación y cálculos son muy claras y sencillas, lo que facilita la evaluación de traductores profesionales. Por otro lado, el funcionamiento es razonable, compensado la desventaja de la evaluación humana, que es muy subjetiva. Cabe mencionar que esta métrica puede desempeñar su papel muy perfectamente en el caso de evaluar un texto de palabras limitadas, pero, no es ideal aplicar esta métrica en un corpus de montón de palabras. Según la figura 12 y la figura 13, se encuentra que el trabajo de encontrar errores, hacer penalización y hacer cálculo cuesta mucho tiempo y mucha energía.

### **Figura 13**

*Registro de errores según SAJ2450 (Sánchez, 2020)*

Categoría de error	Total errores graves (multiplicados por el peso numérico de penalización)	Total errores poco importantes (multiplicados por el peso numérico de penalización)	Puntuación ponderada
terminología	$A = \text{total errores} \times 4$	$B = \text{total errores} \times 2$	$A + B$
sintaxis	$A = \text{total errores} \times 4$	$B = \text{total errores} \times 2$	$A + B$
omisión	$A = \text{total errores} \times 4$	$B = \text{total errores} \times 2$	$A + B$
orden de palabras y concordancia	$A = \text{total errores} \times 4$	$B = \text{total errores} \times 2$	$A + B$
ortografía	$A = \text{total errores} \times 3$	$B = \text{total errores} \times 1$	$A + B$
puntuación	$A = \text{total errores} \times 2$	$B = \text{total errores} \times 1$	$A + B$
otros	$A = \text{total errores} \times 3$	$B = \text{total errores} \times 1$	$A + B$
Suma de las puntuaciones ponderadas			[indicar aquí la suma de todas las puntuaciones]
Puntuación normalizada del texto traducido = suma de las puntuaciones ponderadas ÷ núm. total de palabras en el texto original			

Tabla 2. Registro de errores según SAEJ2450

#### Figura 14

*Pesos numéricos en la métrica SAEJ2450 (Sánchez, 2020)*

Categoría de error	Penalización por error grave	Penalización por error poco importante
terminología	5	2
sintaxis	4	2
omisión	4	2
orden de palabras y concordancia	4	2
ortografía	3	1
puntuación	2	1
otros	3	1

Tabla 1. Pesos numéricos en la métrica SAEJ2450

En base a las dos figuras, se expone un ejemplo. Si la traducción de un texto es de 1000 palabras, después de la evaluación, se encuentra 1 error grave en terminología, 1 grave en sintaxis, 1 grave en omisión, 1 poco importante en orden de palabras y concordancia, 1 poco importante en ortografía, 1 poco importante en puntuación, y un error en otros. En este caso, el resultado de la calidad de evaluación es  $5+4+4+2+1+1+1=18$ ,  $18 \div 1000=0.018$  en una escala de 0 a 1.

## Modelo TAUS

De tercero, se introduce cómo funciona la métrica TAUS (Translation Automation User Society). Esta métrica se desarrolló en 2005 y de acuerdo con Ke Hu, Patrick Cadwell (2006), “Rather than differentiating between guidelines for light and full post-editing, the Translation Automation User Society (TAUS) differentiated between two levels of expected quality, including “good enough” quality, and “human translation quality” (TAUS, 2016).”<sup>69</sup> Es decir, una característica de TAUS consiste en diferenciar dos niveles de calidad esperar.

**Tabla 8**

*Tipología de errores y estudio de las directrices sobre postedición ligera y completa (Establecido propio en base a Vicent Briva-Iglesias (2021), y, Ke Hu, Patrick Cadwell (2006))*

	Errores	PE ligera	PE completa
Precisión (PRE)	1) Interpretación incorrecta del texto en LO, traducción errónea; 2) Melendado del concepto técnico; 3) Omisión (elemento esencial del texto en LO que falta en la traducción; 4) Adición (elementos innecesarios en la traducción, no presentes en el texto en LO); 5) Un segmento con 100% de coincidencia no se ha traducido bien o no es apropiado en el contexto; 6) Texto no traducido;	transmitir el mismo significado del texto original al texto destino	transmitir el mismo significado del TO al TD
Terminología	1) Incumplimiento de la terminología del cliente; 2) Incumplimiento de la terminología del producto; 3) Terminología inconsistente;		ser correcto los términos claves
Lengua (LENG)	Gramática: Incumplimiento de las normas de la LM	no perfecto	perfecto
	Semántica	correcto	correcto
	Ortografía: Errores, uso incorrecto	aplicar	correcto

<sup>69</sup> Hu, K., & Cadwell, P. (2016). A comparative study of post-editing guidelines. In Proceedings of the 19th annual conference of the European association for machine translation (pp. 34206-353).



	de acentos y mayúsculas	normas básicas	
	Sintáxis	ser inusual	ser inusual
Estilo (EST)	1) Incumplimiento de las guías de estilo de la empresa; 2) Inconsistencia con otro material de referencia; 3) Inconsistencia dentro del texto; 4) Traducción literal; 5) Sintaxis adecuada; 6) Uso no idiomático de la LM; 7) Malinterpretación del tono;		no hace falta
Reestructuración			en caso de necesidad
Culturas			en caso de necesidad
información		completo	completo
Otros			

Entonces, según la tabla 8, se encuentra que la métrica TAUS hacer la evaluación por medio de comentar la precisión, terminología, gramática, semántica, ortografía, sintaxis, estilo, cultura, información, etc. Según Susana Cruces Colado (2001), el error de traducción se puede clasificar por “tipo gramatical, de nominalidad léxica, congruencia semántica o de conformidad al conocimiento del mundo y de la experiencia acumulada” (Colado, 2001).<sup>70</sup>

En la parte de introducción de la métrica SAEJ2450, se planea una duda sobre cómo juzgar que un error es grave, o poco importante. En cuanto a esta pregunta, Sánchez (2020) organiza una tabla sobre la identificación del valor de errores de traducción.

**Tabla 9**

*Criterios de TAUS para la asignación de valores a la tipología de errores (Sánchez, 2020)*

Grave	Errores que puedan tener implicaciones en la salud o la seguridad de las personas, o efectos negativos en cuestiones legales o financieras. También se marcan aquellos que puedan violar las pautas de uso geopolítico, dañar la reputación de una empresa, hacer que una aplicación de software se bloquee o que un producto/servicio se
-------	---

<sup>70</sup> Colado, S. C., Real, E., Jiménez, D., Pujante, D., & Cortijo, A. (2001). El origen de los errores en traducción. *Écrire, traduire et représenter la fête*, 813-822.

	modifique. En esta categoría se marcan también los errores que hacen que el texto resulte ofensivo.
Importante	Errores que puedan confundir o inducir a error al usuario u obstaculizar el uso correcto del producto/servicio debido a un cambio significativo en el significado; errores que aparecen en una parte visible o importante del contenido.
Poco importante	errores que no conducen a la pérdida de significado y no confundirán ni engañarán al usuario, aunque sí disminuyen la calidad estilística, la fluidez o la claridad del contenido.
Neutro	Esta categoría sirve para registrar información adicional, problemas o cambios que se deben realizar, pero que no cuentan como errores. Por ejemplo, reflejan la elección léxica que pueda hacer un revisor o una preferencia de estilo. Se indican también errores que se repiten o los cambios en los glosarios que aún no se han implementado u otro tipo de cambios de los que no se ha informado aún al traductor.
Excelente	Esta es una categoría opcional que se puede utilizar cuando se quiere hacer una mención especial a algún problema de traducción resultado de manera excepcional. Al contrario que las otras categorías, esta o supone una penalización.

Sánchez (2020) describe el proceso de hacer cálculo de manera muy clara. En base a esto, se realizarán la evaluación de TA por iFLYTEK en el siguiente capítulo. Antes que nada, es imprescindible aclarar las siglas de los parámetros en las fórmulas para calcular. TPC es Total de Penalización por Categoría;  $T_{EPI}$  es Total de Errores Poco Importantes;  $VP_{WPI}$  es Valor de Penalización de Errores Poco Importantes;  $T_{EI}$  es Total de errores importantes;  $VP_{EI}$  es Valor de Penalización de Errores Importantes;  $T_{EG}$  es Total de errores Graves;  $VP_{EG}$  es Valor de Penalización de Errores Graves;  $T_{ME}$  es Total de Menciones Excelentes;  $V_{ME}$  es Valor de las Menciones Excelentes; TPO es Total de Palabras del Texto Original; PPP es el Porcentaje de Penalización por Palabra.

El proceso de evaluación con la métrica TAUS consiste en (Resumido en base a Sánchez, 2020):

- 1) El traductor es responsable comentar y encontrar los errores de la traducción automática.
- 2) Corresponde el nivel de gravedad de cada error: grave, importante, poco importante, neutro, excelente. Es decir, cuenta el número de  $T_{EPI}$ ,  $T_{EI}$ ,  $T_{EG}$ ,  $T_{ME}$ .
- 3) Hace cálculo del Total de Penalizaciones por Categoría (TPC).

De primero, la penalización de cada error es:

$$VP_{EPI}=0.25$$

$$VP_{EI}=0.5 \text{ (No se regula Sánchez (2020), así se adjunta desde mi punto de vista)}$$

$$VP_{EG}=0.75$$

$$V_{ME}=1$$

La fórmula para calcular el total es :

$$TPC=(T_{EG}\times VP_{EG})+(T_{EI}\times VP_{EI})+(T_{EPI}\times VP_{EPI})-(T_{ME}\times V_{ME})$$

- 4) Hace cálculo del porcentaje del Total de Penalizaciones (TPC) ocupado en el Total de

Palabras del Texto Original (TPO).

$$PPP = (TPC \div TPO) \times 100$$

5) Por último, hace la evaluación de la calidad de la traducción.

$$\text{Calidad de la traducción} = 100 - PPP$$

En resumen, este proceso es muy lógico y sencillo. Se expone un ejemplo: en una traducción de palabras 100, el traductor encuentra 1 error grave, 2 error importante, 3 error poco importante y 1 excelente. Siguiendo el proceso, el total de penalizaciones por categoría es  $(0.75 \times 1) + (0.5 \times 2) + (0.25 \times 3) - (1 \times 1) = 1.5$ . Entonces, el porcentaje de penalización de palabras es  $(1.5 \div 100) \times 100 = 1.5$ . Por fin, se obtiene el resultado de la calidad de la traducción:  $100 - 1.5 = 98.5$

Después de analizar el proceso, el estándar de calificación, se propone una duda pequeña. Desde mi punto de vista, es mejor separar la penalización por categoría grave, importante, poco importante de categoría excelente en el proceso de cálculo. Es que, en caso de incluir la penalización de categoría excelente, el resultado de la evaluación de traducción automática será un poquito confuso. Por un lado, en cuanto a la categoría excelente, no se penalizará sino se premiará en la fórmula para calcular el total de penalización. Así trae un problema, si justo el valor de otros errores penalizados se equivale con el valor de la categoría excelente, la calidad será 100. En este caso, no se puede reflejar la calidad de manera muy clara. Por otro lado, si se separan las dos categorías en la penalización, al final, en el resultado se reflejan una nota de penalización de errores y una nota de menciones excelentes. Por lo tanto, el cliente puede ver muy claramente las limitaciones y ventajas de la traducción automática.

Por fin, luego de investigar esta métrica, se aplicará en el análisis de evaluación en el siguiente apartado. Pero se ajusta la manera para comentar la calidad a través de hacer cálculo de penalización y menciones excelentes por separado.

## **2.4 Particularidad de traductor iFLYTEK**

En este apartado se introduce la empresa iFLYTEK, las cuatro funciones principales, la aplicación del traductor iFLYTEK ahora en China. Hasta ahora, a nivel mundial existen muchas plataformas, herramientas como lo reseñado en el capítulo 1. Sin embargo, la competencia de este traductor consiste en traducción en campos específicos.

La empresa iFLYTEK se estableció en 1999, y ahora es conocida como una empresa líder en el campo de voz inteligente e inteligencia artificial. Esta entidad está dedicada a la investigación de tecnologías como el habla inteligente, la comprensión del lenguaje natural y la visión por computadora. Además, toma como objetivo promover la implementación de inteligencia artificial y aplicaciones industriales a fin de constituir un mundo mejor. Y

en 2006, iFLYTEK lanzó su nuevo producto de serie de traductor. (Sitio web de iFLYTEK, 2023)

Hasta ahora, existen dos series de traductores iFLYTEK, una es máquina de traducción, el otro es bolígrafo de traducción. Y en esta tesis, se utilizará la máquina de traducción, que se dividen en diferentes estilos, como iFLYTEK AI P20 Plus, iFLYTEK AI P20, y iFLYTEK serie S, etc. En base a su corpus de datos de industria de traducción y conocimiento de experto, este traductor está orientado al ámbito más profesional y más específico.

Generalmente, este traductor tiene cuatro funciones principales para satisfacer la necesidad de clientes como viajar, tomar clases, hacer conferencias, realizar estudios, etc. En primer lugar, esta máquina permite realizar traducción e interpretación en tiempo real, admitiendo la traducción instantánea en línea entre 83 idiomas y chino, como inglés, castellano, catalán, francés, etc. En segundo lugar, el traductor de último estilo puede llevar al cabo la traducción e interpretación offline, admitiendo la traducción en tiempo real sin conexión de internet entre 15 idiomas y chino, como, por ejemplo: inglés, japonés, francés, español, chino, etc. En tercer lugar, este producto es capaz de realizar el reconocimiento de acento como: inglés (Reino Unido), inglés (Canadá), inglés (India), inglés (Nueva Zelanda), inglés (Australia), francés (Canadá), bengalí (India), árabe (Irak), árabe (Emiratos Árabes Unidos), árabe (Egipto), español (Argentina), español (México), español (Chile), español (Colombia), portugués (Brasil). Al mismo tiempo, se puede reconocer dialectos chinos como cantonés, sichuanés, dialecto de provincia HeNan, dialecto de Noreste, dialecto de Shandong. Por último, se equipa de terminologías profesionales y específicos en los siguientes campos: comercio exterior, energía, derecho, medicina, deportes, finanzas, informática, energía eléctrica, tecnología de información, ingeniería química, ciencia de materiales, industria mecánica, industria del automóvil, industria de la construcción naval, ingeniería vial, literatura y medio de comunicación.

Excepto de las cuatro funciones principales, con este traductor se puede traducir fotos, conferencias, conversaciones y grabaciones. Según la introducción, la función es muy poderoso. Actualmente, todavía no hay muchos estudios e investigación en cuanto a la calidad de traducción e interpretación en campos especializados. La mayoría de sus usuarios consisten en las personas que viajan a extranjeros o estudian afuera.

En resumen, la particularidad de esta máquina de traducción radica en sus funciones. En cuanto al par de lengua español-chino, los hispanohablantes de diferentes países o diferentes lugares tienen sus acentos, lo que dificulta la traducción e interpretación. Así la función de reconocimiento de acento contribuye a eliminar la barrera, haciendo más fácil mantener la comunicación entre dos partes de interlocutores. Al respecto a la traducción, como lo explicado antes, el sistema consta de terminologías y conocimientos en ciertos campos, incluida la medicina. Entonces, en comparación con los traductores generales, como translate Google, DeepL, etc., en la traducción con iFLYTEK, se puede elevar la calidad con la competencia de corpus relativo. Sin embargo, cabe mencionar que el uso

de iFLYTEK no es muy amplio entre los ciudadanos chinos, sino se usa mucho en la empresa y gobiernos. La causa principal radica en su precio alto. De todas maneras, la empresa IFLYTEK está desarrollándose de manera muy rápida.

### **3. Muestra de datos y metodología**

#### **3.1 Descripción del estudio**

##### **1) Creación de corpus:**

Antes que nada, el corpus consiste en un conjunto lo más extenso y ordenado posible de datos o textos científicos, literarios, etc. que pueden servir de base a una investigación. (Diccionario: Real Academia Española, 2023) En otras palabras, en un corpus se recopila datos o textos en un ámbito específico a fin de hacer un estudio o investigación.

Según Julio Alexander Bernal Chávez y Diana Alejandra Hincapié Moreno (2018), un corpus necesita satisfacer los siguientes requisitos:

**De primero**, “los corpus son porciones de lenguas y de variedades lingüísticas capaces de representar sus tendencias o características.” (Chávez y Moreno, 2018). Entonces, como el enfoque de esta tesis consiste en la traducción de español a chino y después, hacer evaluación de la traducción automática, necesito establecer un corpus bilingüe, español para traducir y chino para servir de referencia en la evaluación automática. Y los dos necesitan corresponderse completamente.

**De segundo**, los datos y textos de un corpus deben ser reales y confiables. Entonces, en el proceso de establecer un corpus, es necesario juzgar la fuente si es oficial o confiable. Se trata de una garantía de autenticidad y fiabilidad del resultado realizado por este estudio.

**De tercero**, los materiales de un corpus pueden de diferentes formas, como orales, textuales, visuales o multimodales. En este estudio, se creará un corpus solo de materiales textuales, es que el tema es traducción. Pero esta metodología también se puede adaptar a un estudio de interpretación, en el que se aplicará corpus orales, visuales o multimodales.

**De cuatro**, hace falta confirmar el tamaño del corpus. Para realizar una investigación, generalmente se prefiere un corpus con montón de palabras para garantizar el resultado más objetivo. Pero teniendo en cuenta la evaluación humana de traducción humana, la cantidad de palabras no puede ser demasiado grande. Así, prefiero elegir establecer corpus de 2000-3000.

**De cinco**, un corpus debe ser representativo. Este estudio se centra en el ámbito sanitario, así todos los materiales en un corpus deben tener que ver con la medicina, la sanidad de manera estrecha. En consecuencia, elijo el tema de tuberculosis, y los textos en el corpus están relacionados con la enfermedad, el tratamiento, las pruebas, la infección y la prevención, etc.

**Otros**, un corpus ha de ser digital, fácil de acceso, y equilibrado.

##### **2) Traducción de corpus:**

Después de establecer el corpus español, se traducirá con la máquina de traducción iFLYTEK. Se adjuntará el corpus original y corpus traducido en el anexo. Con la

traducción, se describirá los problemas surgidos durante el proceso y analizarán las características del texto traducido.

### **3) Evaluación de traducción:**

La evaluación de traducción realizada por IFLYTEK consiste en un punto clave de este estudio. Es que, con la evaluación, se analizará la calidad de TA con la métrica automática BLEU y por medio de la evaluación humana.

Por un lado, en la evaluación automática, se presentará el resultado de BLEU después de terminar el análisis. Además, se propondrán los errores aparecidos en la traducción y se clasificarán según las características. Se adjuntarán los datos y tablas formados en BLEU.

Por otro lado, en la evaluación humana, de primero, voy a revisar la traducción para encontrar los errores. Después, se enumeran los errores y se clasificará según las diferentes categoría. Aquí se aplica el modelo TAUS (Translation Automation User Society) y de acuerdo con el segundo capítulo, se realizarán los cálculos.

El motivo de elegir este tema radica en la necesidad de la realidad. En esta época, debido al desarrollo de tecnología, es fácil viajar por todo el mundo en avión, así el movimiento de población entre España y China se ha convertido en una tendencia más común. Ni siquiera cada año una gran cantidad de alumnos van a España a profundizar el estudio y de visitantes van a España a apreciar el paisaje y culturas. Debido a la barrera de idiomas y contextos culturales, es difícil realizar la consulta al médico sin ayuda de traductores. Sin duda, la sanidad consiste en un tema esencial para la humanidad. Está relacionado estrechamente con la felicidad y la vida de los humanos. Pero no es ideal depender de traductores para solucionar este dilema. No hay suficientes traductores calificados para satisfacer la demanda por las dos dificultades en España y China: falta de formación y falta de código ético para seguir.

El objetivo de este estudio consiste en analizar la capacidad y competencia de traducción de iFLYTEK en el ámbito sanitario. Elijo el corpus limitado al campo de tuberculosis como ejemplo, es que este campo es muy significativo para toda la humanidad. Hasta ahora todavía existen muchos lugares, especialmente, los países menos avanzados, que están corriendo el riesgo de la infección de tuberculosis. Por un lado, tengo intención de promover la información bilingüe para que los ciudadanos de dos países puedan conocer el sistema y noticias de tuberculosis por medio de leer el corpus establecido a fin de saber los síntomas, la influencia y la prevención de esta enfermedad. Por otro lado, excepto de las características de iFLYTEK, quiero analizar los rasgos de evaluación automática y evaluación humana de traducción automática para comparar las ventajas y desventajas a fin de mejorar los modelos y métricas.

## **3.2 Muestra de datos**

### **3.2.1 Creación de corpus**

Después de realizar la búsqueda, en China y España todavía no existe un corpus bilingüe español-chino, sino hay corpus bilingüe inglés-chino e inglés-español. En este caso, para llevar al cabo este estudio, hace falta establecer un corpus bilingüe español-chino en el campo de tuberculosis en especial. Los textos de dos idiomas son paralelos y correspondientes para que se tomen los chinos de referencia en el proceso de evaluación.

Por lo consiguiente, hace falta buscar los textos en español y chino en el campo de la enfermedad de tuberculosis en los sitios web de organizaciones a nivel mundial y instiutos médicos, tales como la Organización Mundial de la Salud, UNICEF, Comité Internacional de la Cruz Roja, New York Department of Health, Centers for Disease Control and Prevention, County of los Angeles Public Health, Ministerio de Sanidad.

En la clase de traducción especializada en el ámbito sanitario, Natalia Maravall (2022)<sup>71</sup> había proporcionado unos recursos electrónicos en cuanto a la enfermedad tuberculosis:

1) Centers for Disease Control and Prevention:

<https://www.cdc.gov/thinktesttreattb/resources.html>

2) County of los Angeles Public Health:

<http://ph.lacounty.gov/tb/factsheets.htm>

3) Mass.gov:

<https://www.mass.gov/service-details/tuberculosis>

4) A NSW Govermente website:

<https://www.health.nsw.gov.au/Infectious/tuberculosis/Pages/tuberculosis-factsheets-languages.aspx>

La selección debe seguir los siguientes principios: de primero, hace falta garantizar la credibilidad y autenticad de la fuente. Con respecto a este punto, las organizaciones enumeradas arriba y los sitios web son fieles y de carácter oficial y formal. De segundo, los textos deben tener que ver con la enfermedad de tuberculosis de manera muy estrecha, tales como la causa de tuberculosis, los síntomas, la prevención, la infección, la prueba, la vacuna, etc. Cuánto más profesional, en los textos hay más terminologías específicas. De tercero, en el proceso de creación de corpus, es imprescindible eliminar las fotos, imágenes para ajustar el formato de texto sencillo (extensión. txt). Por último, los textos deben tener versiones bilingües de español-chino. En todas las organizaciones y sitios web los artículos y las noticias están escritos en inglés, así en el proceso de elección, es necesario revisar si hay texto traducido a otros idiomas.

Por lo consiguiente, en base a los principios anteriores, se seleccionan los siguientes textos:

1) ¿Cómo hablo con mi proveedor médica sobre la tuberculosis (TB) inactiva? (如何与我的医疗服务提供者谈论非活动性结核病 (TB) ? )

2) La infección de la TB. Lo que usted necesita saber para mantenerse saludable. (潜伏结核感染. 关于保持健康的须知)

---

<sup>71</sup> Natalia Maravall, (2022), Traducción especializada en el ámbito sanitario, UD-Maravall-TS, PPT,

- 3) Tuberculosis (TB) (结核)
- 4) Programa de NSW sobre Tuberculosis (TB). Información sobre Tuberculosis (TB) para los contactos (NSW 结核病项目. 接触结核病患者后如何应对)
- 5) La enfermedad de la TB. Tome el control de su salud. (活动期结核病. 掌握个人健康的主动权)
- 6) Tuberculosis e infección por el virus de la inmunodeficiencia humana (VIH) (结核病和人类免疫缺陷病毒 (HIV) 感染)
- 7) Prueba cutánea de la tuberculina (结核菌素皮试)
- 8) Instrucciones para la recogida de apsutum (采集脓液的说明)
- 9) Infección tuberculosa latente y tratamiento preventivo (潜伏的结核病感染和预防治疗)
- 10) Vacunación BCG (卡介苗接种)
- 11) Medicación antituberculosa (结核病药物治疗)
- 12) Información para contactos de tuberculosos (结核病接触者的信息)

Entonces, los doce artículos se pueden considerar como la bibliografía de este corpus en español y en chino. Si tiene interés conocer el corpus de más detalles, es aconsejable subir los sitios webs introducidos anterior para consultar estos textos.

En suma, el corpus en español es de 10729 entradas, mientras el corpus en chino es de 10982 entradas. Para los dos corpus, todavía hace falta organizar el formato. Pero lo más importantes es la eliminación de la parte repetida. En los doce textos, aunque en cada artículo el enfoque es diferente, inevitablemente hay muchas palabras repetidas. La creación de corpus se debe seguir a los criterios del diseño de corpus, la cantidad es limitada, la calidad es original y renovada, así como el corpus debe ser simplicidad de codificación y documentación.

Entonces, se simplifica el corpus en español y chino de acuerdo con los síntomas, la definición de la enfermedad, la vacunación, la medicación, la prevención, y Prueba cutánea de la tuberculina. Además, teniendo en cuenta que la traducción del corpus español se realizará evaluación por humanos, la cantidad de las palabras no puede ser demasiada.

En base a los principios, después de eliminar las partes duplicadas, los caracteres ilegibles, corresponder el contenido de corpus bilingüe y organizar el formato de extensión txt, se obtiene un corpus en español y un corpus en chino con la estructura presentada en la tabla 10. Concretamente, el corpus está dividido en la infección de la tuberculosis, los medios de contagiar el virus, el tratamiento de la infección, dos tipos de pruebas para detectar la infección, los posibles efectos secundarios de las medicinas, la manera para juzgar si está contagiado, y la introducción de vacuna BCG. En cuanto a la enfermedad tuberculosis, este corpus abarcan desde la prevención hasta el tratamiento, incluyendo muchas terminologías específicas. Además, todas estas informaciones provienen de los sitios web de instituto formal o gobierno, así son de calidad y se puede confiar.



**Tabla 10***Estructura y palabras de corpus en español y chino sobre la tuberculosis*

Corpus de tuberculosis en español (totalmente, 1082 palabras)	Corpus de tuberculosis en chino (totalmente, 1702 palabras)
1. ¿Qué es la infección de la TB (LTBI)? (145 palabras)	1. 什么是潜伏结核感染(LTBI)? (248 palabras)
2. ¿Cómo se contagian las personas? (73 palabras)	2. 结核病菌如何传染? (111 palabras)
3. ¿Cuál es el tratamiento para la infección de la TB (LTBI)? (73 palabras)	3. 对潜伏结核感染有哪些治疗方法? (123 palabras)
4. ¿Por qué debo tomar medicina si no me siento enfermo(a)? (98 palabras)	4. 如果我并未感到身体不适,为何仍需服药? (169 palabras)
5. Hay dos tipos de pruebas para detectar la infección de tuberculosis. (108 palabras)	5. 结核病感染检测有两种类型。(149 palabras)
6. ¿Qué pasa si el resultado de la prueba de la piel es positivo? (71 palabras)	6. 皮试结果若呈阳性会怎样? (95 palabras)
7. ¿Cuándo debo hacerme la prueba de la piel? (109 palabras)	7. 若存在以下情况,您应接受测试: (177 palabras)
8. ¿Cuáles son algunos posibles efectos secundarios de las medicinas de la enfermedad de la TB? (54 palabras)	8. 潜伏结核感染(LTBI)治疗药物可能会产生哪些副作用? (85 palabras)
9. No me siento enfermo, ¿cómo puedo saber si he sido contagiado? (134 palabras)	9. 我没觉得身体不适,如何判断被感染呢? (206 palabras)
10. ¿Qué es la Vacuna BCG? (112 palabras)	10. 卡介苗疫苗(BCG)是什么? (143 palabras)
11. La BCG y otras vacunas de bacilo/virus vivo. (105 palabras)	11. 卡介苗疫苗和其他活性疫苗。(176 palabras)

Se adjuntan el corpus en español y en chino en el anexo para que se realizan traducción y evaluación. Después de todo, se revisa y organiza este corpus de acuerdo con Chávez y Moreno (2018):

- 1) Un corpus es una muestra de lengua: En este corpus no se incluyen todos los textos sobre la enfermedad de tuberculosis, sino se seleccionan unos aspectos más destacables como el tratamiento, la infección, los síntomas para servir como referencia. Y el corpus en español contiene 1082 palabras, mientras el corpus en chino contiene 1702 palabras. Los dos corpus son correspondientes completamente.
- 2) Las muestras de un corpus son reales: Las informaciones de este corpus se recogen desde New York Department of Health, Centers for Disease Control and Prevention, County of los Angeles Public Health. Las tres fuentes son de naturaleza oficial y formal. Entonces, el contenido del corpus es exacto y preciso.
- 3) Brindan información adicional: Este corpus es de carácter bilingüe, así con el corpus

no solo se puede conocer las informaciones de tuberculosis, sino hacer el análisis y comparación entre dos idiomas. Con datos y teoría correspondientes, se presenta de manera clara las características y estructuras de la frase español y china.

4) Facilitan la extracción de datos homogéneos y cuantificables: Para comprobar esto, se aplica la herramienta AntConc, programas para realizar la separación para que AntConc puedan reconocer los caracteres.

De primero, los dos textos en español y chino están en formato Microsoft Word, así hace falta convertirlos a formato compatible de texto (con extensión .txt), lo que se puede realizar directamente por Microsoft Word.

De segundo, para el texto en chino, se aplica la herramienta de ROST para realizar la segmentación. Es muy fácil de operar y utilizar, según la figura 15, este software es un archivo en nube Baidu. después de descargarlo (<https://pan.baidu.com/s/1TPMhpEOAT--BfNPmRdTzqA>), haga clic en el análisis funcional en la esquina superior derecha y seleccione la segmentación de palabras, y luego elija el archivo para realizar análisis. Así el texto chino se realiza la segmentación. De acuerdo con la figura 16, se puede ver el corpus chino segmentado según el uso, la categoría gramatical y luciones, etc. Para el documento en español no es necesario, es que los dos idiomas tienen diferentes características, que la frase china no se separa de manera automática, así hace falta utilizar programas para realizar la segmentación. Esta herramienta es fácil de operar y tiene muchas funciones. Es imprescindible aplicar en los estudios e investigaciones de chino y otros idiomas extranjeros. Es recomendable probarla con el sitio web y seguir las instrucciones.

**Figura 15**

*Segmentación de palabras chinas con ROST*



**Figura 16**

*Comparación de texto segamentado y no segmantado*

结核病菌如何传染？

当活动期结核病患者咳嗽、打喷嚏、发笑、讲话或唱歌时，结核病菌便会在空气中传播。若您吸入带有结核病菌的空气时，便会

对潜伏结核感染有哪些治疗方法？

若您受结核病菌感染，则会患结核病。您可能会立即发病，也可能于几年之后发病。对潜伏结核感染可采用药物治疗；最常用的

如果我并未感到身体不适，为何仍需服药？

如果您不服用潜伏结核感染治疗药物，这种感染则会发展为活动期结核病。开始服药后，许多结核病菌将被杀灭，但仍会有一些

什么是潜伏结核感染 (LTBI)？

潜伏结核感染或 LTBI 系指您体内呈“休眠”状态的存活结核病菌。潜伏结核感染患者并无任何症状、不会感活动期结核病系指人体内的结核病菌仍然存活并已“苏醒”。活动期结核病患者会有严重咳嗽、夜间盗汗、疲

结核病菌如何传染？

当活动期结核病患者咳嗽、打喷嚏、发笑、讲话或唱歌时，结核病菌便会在空气中传播。若您吸入带有结核病菌

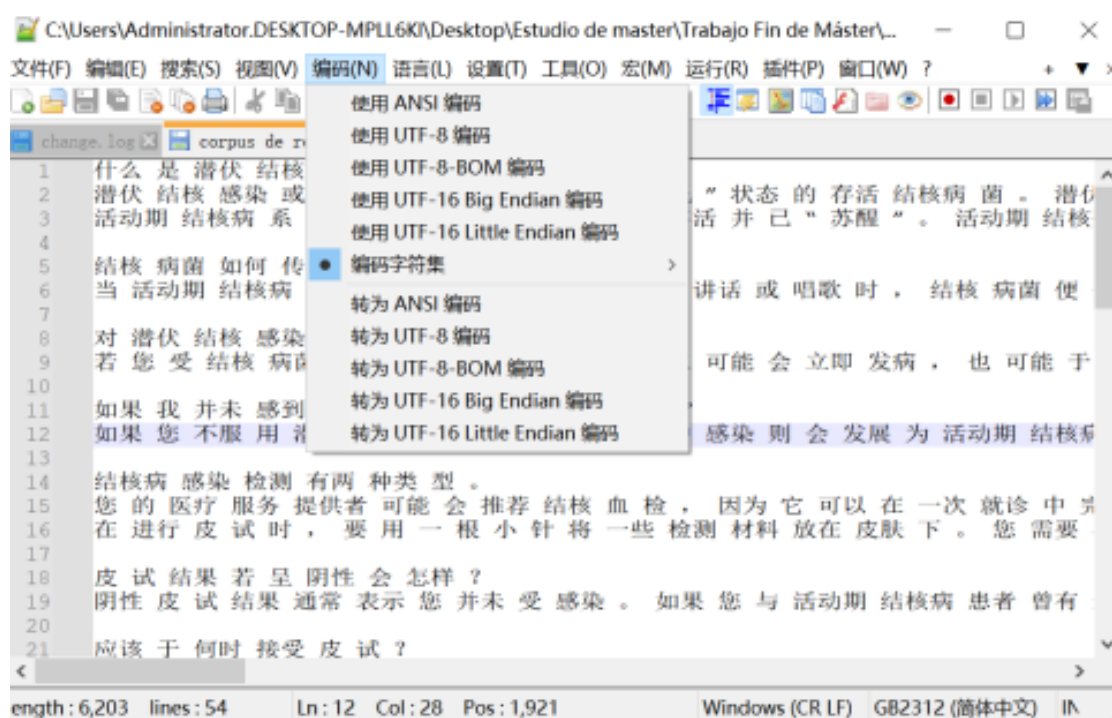
对潜伏结核感染有哪些治疗方法？

若您受结核病菌感染，则会患结核病。您可能会立即发病，也可能于几年之后发病。对潜伏结核感染可采用

De tercero, cabe destacar la importancia transformar la manera de codificación a “UTF-8”. Si no, la herramienta AntConc no puede reconocer y funcionar. Solo con la segmentación, no es suficiente para operar la AntConc. Aquí se utiliza el editor de texto “Notepad ++” para convertir el método de codificación. Según la figura 17, se muestra los pasos de uso. Después de descargar el software (se adjunta el sitio web para descargar: <https://notepad-plus-plus.org/downloads/v8.5.2/>), abre el corpus en español y el en chino, seleccione la función de la configuración de codificación, haga clic para convertir a codificación UTF-8, y guárdelos. De esta manera, se cambia la manera de codificación de los dos textos, para que sea compatible con AntConc. Esta herramienta tiene otras funciones, tales como selección de palabras claves, comparación y análisis de textos, etc.

Figura 17

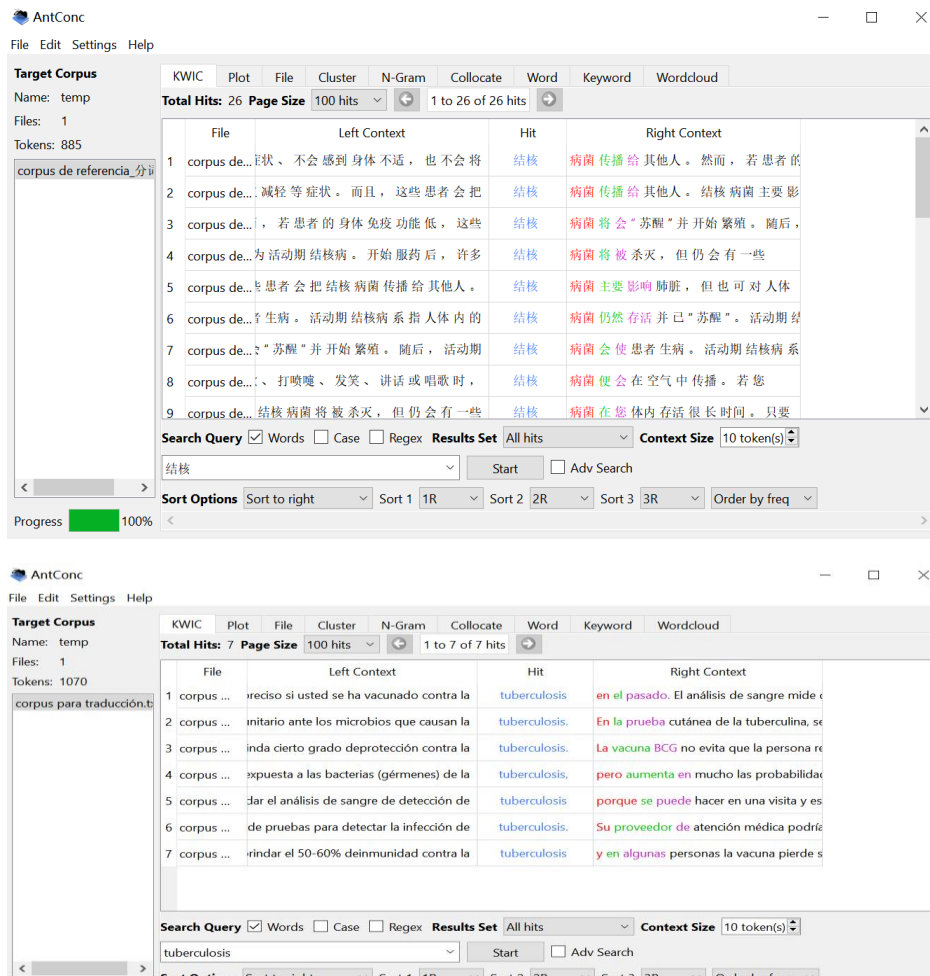
Método de configurar la codificación de Notepad ++



De cuatro, al respecto a AntConc, en el proceso de operación, surge con frecuencia el problema de “UnicodeDecodeError: 'UTF-8' codec can't decode byte 0xca in position 5: invalid continuation byte”, entonces, es importante saber cómo solucionarlo. En base a los tres pasos anteriores, aquí se utiliza AntConc para realizar la gestión del corpus y el análisis textual. Esta herramienta es usada mucho en el estudio y redacción de tesis debido a la función de compilación de archivos de texto en forma de corpus de forma sencilla y eficaz. En primer lugar, abra esta aplicación, haga clic de archivos, seleccione “abrir archivos” para subir el corpus de destino. Después, se puede ver la concordancia, n-gramas, análisis del tipo de texto, palabras clave, lista de palabras, ect. Además, también se puede localizar los términos específicos en el texto para revisar su uso y combinación. Según la figura 18, se ven las concordancias y la localización de “tuberculosis” y “结核”, lo que significa que, por medio de corpus, no solo se puede realizar el análisis de textos, sino también facilitar la extracción de datos para se aplicarán en otros estudios lingüísticos o académicos.

**Figura 18**

*Concordancia de “tuberculosis”, “结核” en el corpus en español y en chino*



Son representativos y diversos: Como se explica lo anterior, en un corpus no se puede recoger toda la información de un tema o un idioma, un corpus debe ser representativo. En cuanto a una enfermedad, la mayoría de los textos de esta enfermedad suelen describir sus síntomas, manera para prevenir y tratar, método de pruebas, las causas, etc. De acuerdo con la tabla 10, se abarcan casi todos los aspectos representativos en el corpus en chino y en español.

5) Con respecto a otros criterios, la recopilación de este corpus también satisfacer los siguientes requisitos: el formato del corpus es digital; el corpus es de fácil acceso; el corpus debe tener el equilibrio en representatividad, variedad y tamaño.

### **3.2.2 Traducción de corpus (español-chino)**

En esta parte, se describe el proceso de realizar la traducción y los rasgos destacables del traductor iFLYTEK. En el segundo capítulo se expone que el traductor IFLYTEK tienen función de traducción e interpretación en tiempo real, reconocimiento de acento, reconocimiento de dialecto y funcionamiento sin conexión de internet. Además, la tecnología iFLYTEK está equipada con terminologías y conocimientos específicos, mientras puede traducir fotos, conferencias, conversaciones, grabaciones, etc. El enfoque del estudio consiste en la traducción. Hay dos maneras para realizar la traducción, una es utilizar la función de traducir las fotos con la máquina de mano, otra es emplear un traductor online diseñado por la empresa.

Entre las dos maneras, prefiero el traductor online diseñado por la empresa iFLYTEK y su sitio web es: <https://www.xfyun.cn/?ch=xfow>. En comparación con el traductor de mano, esta plataforma está encaminada al explotador individual o empresarial, donde se reúnen muchos servicios no limitados a interpretación y traducción. Es más profesional y multifuncional. La ventaja del traductor de mano consiste en la accesibilidad y tamaño portátil. Se puede elegir de acuerdo con la necesidad. A través de esta plataforma, tenemos acceso a utilizar reconocimiento de voz, interacción de voz fuera de línea, síntesis de voz, extensión de voz, traducción automática, procesamiento de lenguaje natural, etc. La función de traducción automática admite interpretación simultánea, traducción de imágenes, traducción de documentos, traducción de páginas web, traducción de audio y video, traducción de texto, etc.

Aquí se utiliza la TA de documento para traducir el corpus organizado referido en el apartado anterior. Como un traductor automático, la eficiencia de traducción es muy alta, que solo tardan unos segundos llevar al cabo traducir un texto en español de 1082 palabras. Se adjunta el resultado en el anexo. Al mismo tiempo, se resumen rasgos de este traductor según la sensación de experiencia: por un lado, es fácil de operar este traductor, que es compatible con 23 formatos, tales como PDF, TEXT, PPT, DOCS, XLS, XLSX, PPTX, etc. No hace falta convertir el formato de documento. Después de terminar la traducción, el resultado sale en formato de solo texto, lo que facilita el análisis o investigación posterior. Por otro lado, con la lista de diferentes funciones enumeradas, el

diseño de la página de web es muy sencillo y de color puro. Con este traductor se puede satisfacer diferentes necesidades en la traducción o interpretación.

### 3.3 Evaluación y Análisis

Esta parte es el enfoque importante de esta tesis. Se presenta el resultado de la evaluación de traducción realizada por iFLYTEK por medio de métrica automática y modelo humano y se analizan los errores con todo detalle.

#### 3.3.1 Evaluación automática con BLEU

Hasta este apartado, están organizado los tres documentos: texto original en chino (de referencia), texto original en español, texto traducido al chino por iFLYTEK (véanse en los anexos). Entre los tres, se toma el texto traducido como objeto de investigar y estudiar.

Con las herramientas reseñadas en la creación de corpus, los tres archivos son de formato de extensión (.txt). Antes que nada, suba el sitio web de Interactive BLEU score evaluador y seleccione los tres textos a opciones correspondientes: paso 0: archivo de origen, paso 1: archivo de traducción humana, paso 2: archivo de traducción automática. En el primero paso, he encontrado un problema, que el sistema BLEU no puede reconocer e interpretar los textos. En cuanto al idioma español, no se muestra acentos en las letras. En cuanto al idioma chino, no se puede distinguir ninguno.

**Figura 19**  
*Porblema de Mojibake de BLEU*



Para solucionarlo, he intentado muchas veces. He eliminado todos los textos y reorganizado. Por fin, se encuentra que se debe a la configuración de codificación. En los textos se había aplicado la codificación de ANSI, pero el sistema BLEU no puede reconocer, así hay que modificarla a la codificación de UTF-8. Es un punto calve a que deben prestar atención, es que es fácil ignorarlo y es difícil solucionar si no había tenido

en cuenta. Aquí se utiliza la herramienta ROST Content Mining (se introduce en el apartado anterior) a fin de configurar la codificación. Después de transformarla, BLEU consigue distinguir las palabras del español y chino.

## Figura 20

### Problema de correspondencia de textos

Sentence	BLEU	Length ratio	Text
15			
Source	-	-	En la prueba cutánea de la tuberculina, se usa una aguja pequeña para poner algo de material de análisis debajo de la piel. Usted deberá volver a su proveedor de atención médica en 2 o 3 días para ver si hay una reacción.
Human	100.00	1.00	应该于何时接受皮试？
Machine	2.21	5.29	在结核菌素皮试中，用小针？在皮肤下放一些分析材料。您必须在2或3天内返回您的医疗保健提供者，看看是否有反应。

La métrica BLEU no puede ajustar la correspondencia de texto de origen, de referencia, y de traducción automática. Por lo tanto, hace falta adaptar las frases de manera manual. Cabe mencionar que es recomendable segmentar los textos oración por oración o párrafo por párrafo. Si no se separa, hay un montón de frases parecidas en el marco de Text sin correspondencia, dificultando la comparación y el análisis en el apartado posterior. Se analizan según la estructura introducida en la tabla 10.

## Figura 21

### Resultado de BLEU de la traducción realizada por iFLYTEK

BLEU:	<b>17.77</b>
Precision x brevity:	18.30 x 97.09
Type	1-gram 2-gram 3-gram 4-gram
Individual	54.52 24.23 12.54 6.78
Cumulative	52.93 35.29 24.75 17.77
Export data	<a href="#">CSV</a>

Según la figura 21, la nota final dado por BLEU es 17.77. Desde de mi punto de vista, esta nota no es el único estándar para evaluar la calidad de la traducción, sino un indicador para mostrar el nivel de similitud entre el texto de referencia y el texto de traducción automática. Por lo siguientes, se analiza concretamente la calidad de traducción por medio de comentar su precisión, terminología, semántica, sintaxis, formas, estilos, estructura, culturas, etc.

### La primera parte:

## Figura 22

## Comparación de la traducción humana y automática de la primera parte

Human	100.00	1.00	什么是潜伏结核感染(LTBI)? 潜伏结核感染或LTBI系指您体内呈“休眠”状态的存活结核病菌。潜伏结核感染患者并无任何症状、不会感到身体不适,也不会将结核病菌传播给其他人。然而,若患者的身体免疫功能低,这些结核病菌将会“苏醒”并开始繁殖。随后,活动期结核病菌会使患者生病。活动期结核病菌系指人体内的结核病菌仍然存活并已“苏醒”。活动期结核病患者会有严重咳嗽、夜间盗汗、疲劳和体重减轻等症状。而且,这些患者会把结核病菌传播给其他人。结核病菌主要影响肺部,但也可对人体其它器官产生影响。若不加以治疗,结核病可导致死亡。
Machine	17.28	0.90	什么是结核病感染? 结核病或LTBI感染意味着结核菌是活的,但在您的体内“沉睡”。如果您感染了结核病,您就不会有任何症状,您可能不会感到不适,也不会将结核病菌传染给其他人。然而,如果您的免疫系统减弱,这些结核病菌可能会“苏醒”并繁殖。然后,您可能会患上活动性结核病。如果您患有结核病,您会出现剧烈咳嗽、夜间出汗、疲劳和体重减轻等症状。此外,它还会将结核病菌传染给他人。结核病特别影响肺部,但也会影响身体的其他部位。如果不治疗,结核病可能会导致死亡。

En la primera parte, el contenido principal es explicar la definición de tuberculosis. La infección de la tuberculosis se divide en dos diferentes tipos. Por una parte, se llama tuberculosis latente, que se refiere a que los gérmenes no están activos sino dormidos. En esta situación, los pacientes no tienen síntomas obvios, ni contagiar el virus de tuberculosis a otros demás. Pero existe una posibilidad de revivir de los gérmenes en caso de la debilitación del sistema inmunológico. Por otra parte, se llama la enfermedad de tuberculosis, que se refiere a que los gérmenes son muy activos. En este caso, los pacientes tienen síntomas evidentes y pueden contagiar los virus a los demás.

Según la figura 22, se ve que la similitud de dos versiones de traducción es de 17.28%. Las palabras marcadas en color azul son iguales o parecidas, mientras que las palabras marcadas en color rojo son diferentes o desiguales. En base a la traducción humana, se organizan las diferencias de traducción y se analizan:

- 1) “潜伏”: Se corresponde a la palabra “latente” de LTBI en el texto original. En la traducción automática, se ignora la traducción de esta palabra. Además, no se traduce la palabra propia “LTBI”, que significa latente tuberculosis. La versión correcta debe ser “潜伏结合感染”. Entonces, en cuanto al primer punto, no está bien la traducción de esta palabra de TA.
- 2) “系指”: Se corresponde a “significa que “en el texto original. En la TA, se traduce como “意味着”, que es una traducción literal de “significa”. La traducción no es falsa, pero no es igualmente adecuada como la de traducción humana.
- 3) “呈”: Se corresponde a “están” para describir un estado de los gérmenes. En la TA, se traduce como “是”. No es muy exacto, porque los gérmenes ahora están en un estado latente, pero pueden revivir en cierto caso. Es un estado temporal sino una naturaleza. Así, no es muy exacto.
- 4) “休眠”: Se corresponde a “dormidos” en el texto original. En la TA, se traduce como “沉睡”. Las dos traducciones no tienen diferencia en el uso y significación. Entonces, la TA está bien en este punto.
- 5) “状态”: No hay palabras para corresponder en el texto original. La traducción humana quiere destacar el estado actual y temporal de los virus. No se traduce en la traducción automática, y no hay efecto al significado de toda la frase. La ignorancia aquí es aceptable.
- 6) “存活”: Se corresponde a “vivos” en el texto original. En la TA, se traduce como “活的”. La TH destaca que el estado de los gérmenes es existente y vivo. Pero la TA no es falsa, solo no tan buena como la TH.



- 7) “患者”: Se corresponde a “ustedes” en el texto original. En la TA, se traduce como “您”. Obviamente es una traducción literal y directa. La TH destaca la relación de médico y pacientes, así es más profesional. Entonces, la TA no es falsa, pero es mejor la TH.
- 8) “无”: Se corresponde a “no tendrá” en el texto original. En la TA, se traduce como “就不会有”. El significa no es falso, pero el uso de “就” no es muy adecuado. Con esta traducción, se destaca que la condición necesaria para no tener síntomas es tener la infección de la tuberculosis. Pero lo que quería expresar el texto original consiste en que, en el caso de la infección de tuberculosis, usted no tiene síntomas. Entonces, la TA no es muy exacto.
- 9) “传播”: Se corresponde a “contagiar” en el texto original. En la TA, se traduce como “传染”. En este caso, la TA es mejor en comparación con la TH. Es que la TH enfatiza el proceso de transmitir el virus, mientras que la TA destaca la posibilidad de infectar a otras personas.
- 10) “若”: Se corresponde a “si” en el texto original. En la TA, se traduce como “如果”. Son completamente iguales y no hay diferencia entre dos traducciones.
- 11) “功能”: Se corresponde a “sistema inmunológico” en el texto original. En la TA, se traduce como “系统”. Aunque la TA es traducción literal, es correcto. Aquí las dos son iguales.
- 12) “低”: Se corresponde a “se debilita” en el texto original. En la TA, se traduce como “减弱”. La TA es más adecuada aquí en comparación con la TA, porque la TA enfatiza el movimiento dinámico del sistema inmunológico. Antes, los gérmenes están dormidos, pero cuando el sistema se convierte débil, los gérmenes pueden despertar. La TH no tiene este sentido.
- 13) “开始”: Se corresponde a “pueden” en el texto original. En la TA, se traduce como “可能会”. La TA es más exacta, es que en el texto original no se describe una acción y un punto para comenzar, así no se puede usar “开始”.
- 14) Se corresponde a “entonces” en el texto original. En la TA, se traduce como “然后”. La TH significa una posibilidad en caso de despertar los gérmenes, mientras que la TA destaca una inevitabilidad o un resultado de enfermarse. Las dos están bien.
- 15) “活动期”: Se corresponde a “TB activa” en el texto original. En la TA, se traduce como “活动性”. La TA destaca una naturaleza de la tuberculosis, mientras que la TH enfatiza que los gérmenes de TB están en un estado o plazo activo. Entonces, la TH es más adecuada que la TA.
- 16) “使患者生病”: Se corresponde a “enfermarse” en el texto original. En la TA, se traduce como “您可能会染上”. Los dos sujetos de la TA y TH son diferentes. En la TA, el sujeto es el paciente, una voz activa, mientras que, en la TH, el sujeto es el virus de tuberculosis. Las dos formas están bien.
- 17) “人体内”: Esta frase “活动期结核病是指人体内的结核病菌人啊果然存活并已经苏醒” es repetido como una frase de arriba. No hay contenido en el texto original para corresponder.
- 18) “严重”: Se corresponde a “fuerte” en el texto original. En la TA, se traduce como “剧烈”. Aquí para describir la gravedad de tos, la TA es mejor.
- 19) “而且”: Se corresponde a “Además, “en el texto original. En la TA, se traduce como “此外”. El significado de los traducción es igual.

20) “把.....传播给”: Se corresponde a “contagiar a “ en el texto original. En la TA, se traduce como “将...传染给”. La traducción automática es mejor. La razón es como el noveno punto.

21) “肺脏”: Se corresponde a “los pulmones” en el texto original. En la TA, se traduce como “肺部”. No hay diferencia entre las dos.

22) “可对人体其他器官产生影响”: Se corresponde a “puede afectar otras partes del cuerpo” en el texto original. En la TA, se traduce como “也会影响身体的其他部位”. La TH enfatiza el efecto a los órganos, mientras que la TA se refiere a partes del cuerpo, que es más general. No existen gran diferencia, todo bien.

23) “若不加以治疗”: Se corresponde a “Si no se trata” en el texto original. En la TA, se traduce como “如果不治疗”. El sujeto de TH es la enfermedad, mientras que el sujeto de TA es paciente. Las dos son buenas.

24) “可”: Se corresponde a “puede” en el texto original. En la TA, se traduce como “可能”. Son completamente iguales.

En resumen, entre 24 diferencias de traducción humana y traducción automática,

- a. La TA es mejor: (9), (12), (14), (18), (19), (17)
- b. La TH es mejor: (1), (2), (3), (6), (7), (8), (15);
- c. La TA y la TH ambas son adecuadas: (4), (5), (10), (11), (16), (19), (21), (22), (23), (24);
- d. La TA ni la TH son adecuada: 0

Entonces, entre 24 diferencias, el número de los casos de que la TA es mejor es 7; el de los casos de que la TH es mejor es 7; el de los casos de que ambas son adecuadas es 10.

### La segunda parte:

En la segunda parte, el contenido principal es la manera de contagiarse las personas. Por un lado, es posible infectarse si respira el aire con gérmenes, los que se propagan en el aire cuando un paciente con tuberculosis habla, canta, tose, etc. Por otro lado, no se puede infectarse solo por medio de los artículos diarios o el contacto físico.

**Figura 23**

*Comparación de la traducción humana y automática de la segunda parte*

Sentence 2	BLEU	Length ratio	Text
Human	100.00	1.00	结核 病菌 如何 传染? 当 活动期 结核病 患者 咳嗽、打喷嚏、发笑、讲话 或 唱歌 时, 结核 病菌 便 会 在 空气 中 传播。若 您 吸入 带有 结核 病菌 的 空气 时, 便 会 受 感染。然而, 您 不会 经由 衣物、盘子、食物、床上 用品 或 与 活动期 结核病 患者 的 实际 接触 而 患 结核病。
Machine	24.59	1.07	人 是 怎么 传染 的? 结核 病菌 在 空气 中 传播 时, 患有 活动性 结核病 的 人 咳嗽, 打喷嚏, 大笑, 说话 或者 唱歌。如果 您 呼吸 的 空气 中 有 结核 病菌, 您 可能 会 被 感染。然而, 您 不能 通过 衣服、盘子、食物、床上 用品 或 与 患有 活动性 结核病 的 人 的 身体 接触 而 感染 结核病。

Según la figura 23, la tasa de similitud de traducción humana y traducción automática es de 23.59 %. En base a la TH, se comparan y se analizan las zonas marcadas en color rojo:

1) “结核病菌如何传染”: Se corresponde a “Cómo se contagian las personas” en el texto

original. En TA, se traduce como “人是怎么传染的”. Aquí la traducción automática es falsa, es que se equivoca la voz activa y pasiva. La persona está infectada sino es el sujeto de contagiar. La TH es mejor que la TA.

2) “发笑”: Se corresponde a “se ríe” en el texto original. En TA, se traduce como “大笑”. La TH contiene un significado “laugh”, mientras que la TA contiene “laugh” y “smile”, así como en el texto original no se define muy claramente, la TH es más adecuada.

3) “讲话”: Se corresponde a “habla” en el texto original. En TA, se traduce como “说话”. Las dos son de dignificado completamente igual.

4) “便”: No hay contenido correspondiente. Se expresa una estructura de condición, “当患者讲话时, 细菌便会传播”, “cuando el paciente habla, el virus se propaga”. La estructura de TA es “病菌传播时, 病人讲话”, “cuando el virus se propaga, el paciente habla”. Entonces, la estructura de TA es falsa.

5) “吸入”: Se corresponde a “respira” en el texto original. En TA, se traduce como “呼吸”. La estructura de las frases traducidas es diferente. La traducción humana es “若您吸入带有结核病菌的空气时” con la estructura “si respira el aire que contiene los gérmenes”, mientras que la traducción automática es “如果您呼吸的空气中有结核病病菌时”, con la estructura “si el aire que respira contiene los gérmenes”. En comparación con la TH, la TA es más ajustada al texto original.

6) “带有”: Se corresponde a “tiene” en el texto original. En TA, se traduce como “有”. Las dos son de significado mismo.

7) “结核”: Se corresponde a “la TB” en el texto original. En TA, se traduce como “结核病”. En el texto original, “la TB” se combina con “gérmenes”. En este contexto, la TB puede ser enfermedad o todavía no se desarrolla a una enfermedad. Entonces, la TH es más adecuada.

8) “便”: No hay contenido correspondiente en el texto original. Es una estructura de frase condicional “si...” traducido en chino “若...便...” en la TH, “如果..会” en la TA. Los dos son estructuras de condición usadas en chino.

9) “受”: Se corresponde a “infectarse” en el texto original. Es una expresión de voz pasiva, igual como la TH, “被...”. Los dos tiene misma función.

10) “不会”: Se corresponde a “puede” en el texto original. En TA, se traduce como “不能”. Aquí se destaca una posibilidad sino una capacidad, así TA no es muy adecuada.

11) “经由”: Se corresponde a “a través de” en el texto original. En TA se traduce como “通过”. Son sinónimos, así las dos son adecuadas.

12) “衣物”: Se corresponde a “ropa” en el tetxo original. En TA se traduce como “衣服”. Son sinónimos, así las dos son adecuadas.

13) “实际”: Se corresponde a “físico” en el texto original. En TA se traduce como “身体”. La TH enfatiza el contacto directo con una persona, mientras que la TA destaca el contacto con el cuerpo de una persona. Como la palabra “físico” contiene estos dos significados, las dos ambas son adecuadas.

En resumen, entre 13 diferencias de traducción humana y traducción automática,

a.La TA es mejor: (5);

b.La TH es mejor: (1), (2), (4), (7), (10);

c.La TA y la TH ambas son adecuadas:(3), (6), (8), (9), (11), (12), (13);

d.La TA ni la TH son adecuada: 0

Entonces, entre 13 diferencias, el número de los casos de que la TA es mejor es 1; el de los casos de que la TH es mejor es 5; el de los casos de que ambas son adecuadas es 7.

### La tercera parte:

En la tercera parte, el contenido principal consiste en el tratamiento para la infección. Después de infectarse por la tuberculosis, puede enfermarse inmediatamente o varios años después. Puede tomar medicamento de isoniazida (INH) para tratar la LTBI durante seis hasta nueve meses.

### Figura 24

*Comparación de la traducción humana y automática de la tercera parte*

Sentence 3	BLEU	Length ratio	Text
Human	100.00	1.00	对潜伏结核感染有哪些治疗方法?若您受结核病菌感染,则会患结核病。您可能会立即发病,也可能于几年之后发病。对潜伏结核感染可采用药物治疗;最常用的药物为雷米封(INH)。服药期必须持续六至九个月。(儿童及受艾滋病毒感染(HIV)者需要更长的服药期;至少应为九个月。)
Machine	7.91	0.91	结核病感染的治疗方法是什么?如果您感染了结核病菌,您可能会患上结核病。这可能会立即发生,也可能发生几次?你们等一下。LTBI可以用药物治疗。最常见的是异烟肼(INH)。这种药应该服用六到九个月。还有一种较短的方案,结合每周服用12周的两种药物。

Según la figura 24, se encuentra que el porcentaje de similitud entre la TA y TH es de 7.91 %. Se comparan y se analizan las diferencias marcadas en color rojo:

- 1) “有哪些治疗方法”: Se corresponde a “cuál es el tratamiento” en el texto original. En TA se traduce como “治疗方法是什么”, que es una manera de traducción literal, pero las dos tienen mismo significado.
- 2) “受感染”: Se corresponde a “se infectó” en el texto original. En TA se traduce como “感染病菌”. La TH es más formal, de voz pasiva, mientras que la TA es más coloquial, ignorancia de voz pasiva. Entonces, en un programa para propagar la información médica, debe ser seria y exacta. La TH es más adecuada aquí.
- 3) “发病”: Se corresponde a “enfermarse...suceder”. En TA se traduce como “发生”. La TH organiza las dos frases “puede enfermarse, puede suceder”, así es más concisa y precisa. La traducción automática de esta frase tiene ambigüedad. Con esta traducción se puede referirse a que el asunto de adolecer la enfermedad o ataque de la enfermedad puede suceder. Entonces, aquí la TH es más exacta.
- 4) “几年之后发病”: Se corresponde a “varios años después”. En TA se traduce como “发生几次”, que es completamente falsa.
- 5) “潜伏结核感染”: Se corresponde a “LTBI”. En TA se traduce como “LTBI”, que debería traducirse las Siglas. El nombre completo es latente tuberculosis. Además, una frase de “你们等一下” en TA también es falsa, que no hay contenido para corresponder.
- 6) “常用的”: Se corresponde a “lo más usual” en el texto original. En TA se traduce como “常见的”, la cual significa “lo más visto”, así no es muy exacta.
- 7) “雷米封”: Se corresponde a “isoniazida (INH) en el texto original. En TA se traduce como “异烟肼”. Son la misma medicina para tratar la tuberculosis.
- 8) “服药期”: Se corresponde a “Debe tomarse esta medicina durante” en el texto original.

En TA se traduce como “这种药应该服用”. La estructura de TH es de subjuntivo, que es más concisa, mientras que la de TA es verbo más objeto, que es más clara. Entonces, las dos ambas son adecuadas.

9) “必须”: Se corresponde a “debe” en el texto original. En TA se traduce como “应该”. La TH destaca la obligatoriedad, y esto depende del tono del autor. Ambas son buenas.

10) “儿童及受艾滋病感染 (HIV) 需要更长的服药期; 至少应该九个月”: No hay contenido para corresponder, que es una laguna. En TA se traduce esta frase como “还有一种较短的方案, 结合每周服用 12 周的两种药物。” Es una traducción correcta .

En resumen, entre 10 diferencias de traducción humana y traducción automática,

a.La TA es mejor: (10);

b.La TH es mejor: (2), (3), (4), (5), (6);

c.La TA y la TH ambas son adecuadas:(1), (6), (8), (9);

d.La TA ni la TH son adecuada: 0

Entonces, entre 10 diferencias, el número de los casos de que la TA es mejor es 1; el de los casos de que la TH es mejor es 5; el de los casos de que ambas son adecuadas es 4.

#### La cuarta parte:

En la cuarta parte, el contenido principal es explicar la necesidad de tomar medicamentos en caso de no sentirse mal. Por un lado, sin tratamiento con medicina, se puede desarrollar la enfermedad la infección de TB. Por otro lado, hace falta insistir tomar medicina para evitar que los gérmenes despierten o se multipliquen. Es importante recibir el tratamiento según las instrucciones del personal médico.

#### Figura 24

*Comparación de la traducción humana y automática de la cuarta parte*

Sentence 4	BLEU	Length ratio	Text
Human	100.00	1.00	如果我并未感到身体不适, 为何仍需服药? 如果您不服用潜伏结核感染治疗药物, 这种感染则会发展为活动期结核病。开始服药后, 许多结核病菌将被杀灭, 但仍会有一些结核病菌在您体内存活很长时间。只要人体内有结核病菌存活, 这些病菌便会“苏醒”、繁殖并使您生病。若您服用结核病治疗药物, 则必须按照医生或医疗保健服务提供者所开的处方服用全部药物; 这十分重要。
Machine	19.69	0.85	如果我不觉得不舒服, 为什么要吃药? 如果您不服用治疗结核病的药物, 您可能会患上结核病。许多结核病菌在开始服药后死亡, 但也有一些在体内存活了很长时间。只要您体内还有结核病菌, 它们就会“苏醒”, 繁殖并使您生病。如果您正在服用结核病药物, 重要的是按照医生或卫生专业人员的指示完成您的药物治疗。

Según la figura 24, la tasa de la similitud de traducción humana y traducción automática es de 19.69%. De acuerdo con las palabras marcadas en rojo, se comparan y se analizan las diferencias entre las dos versiones traducidas:

1) “并未”: Se corresponde a “no” en el texto original. En TA se traduce como “不”. Son sinónimos y no tienen diferencia.

2) “感到身体不适”: Se corresponde a “me siento enfermo(a)” en el texto original. En TA se traduce como “觉得不舒服”. La TH es de estilo escrito, mientras que la TA es de estilo más coloquial. No tienen diferencia en el significado, sino en la forma de

expresión.

3) “为何仍需服药”: Se corresponde a “Por qué debo tomar medicina” en el texto original. En TA se traduce como “为什么要吃药”, que es una traducción literal. Prefiero la traducción humana.

4) “这种感染会发展”: Se corresponde a “Puede desarrollar” en el texto original. En TA se traduce como “您可能患上”, con sujeto de usted. En la TH, el sujeto es la infección. Las estructuras ambas son adecuadas.

5) “将被杀灭”: Se corresponde a “Muchos gérmenes de la TB mueren” en el texto original. En TA se traduce como “死亡”, que es más literal en comparación con TH. Prefiero la TH, es que destaca el efecto de medicina para los virus.

6) “仍”: Se corresponde a “permanecen” en el texto original. En TA se traduce como “也”. Las dos son sinónimos, una es de estilo escrito y otra es estilo coloquial, pero ambas son adecuadas aquí.

7) “人体内”: Se corresponde “en el cuerpo” en el texto original. En TA se traduce como “您体内”, de la segunda persona, usted, mientras que la TH es de la tercera persona. ambas traducciones son buenas.

8) “必须”: No hay contenido para corresponder. Es una expresión de tono a destacar la importancia de seguir las instrucciones del médico. En TA no se interpreta, pero no influencia el significado de la frase. Entonces, las dos son de mismo efecto.

9) “医疗保健服务提供者”: Se corresponde a “el médico o profesional de la salud” en el texto original. En TA se aplica una técnica de traducción literal y se traduce como“卫生专业人员”. No hay diferencia en el significado entre las dos.

10) “处方”: Se corresponde a “lo haya indicado” en el texto original. En TA se traduce como“指示”. La TH destaca las inscripciones dadas por el proveedor del servicio sanitario, mientras que la TA enfatiza las instrucciones de los médicos. Aquí la TA es más adecuada, es que lo indicado no solo se limita a registros sino también puede ser consejos o notas del personal sanitario.

11) “全部药物”: Se corresponde a “termine su medicina” en el texto original. En TA se traduce como “完成您的药物治疗”, que es una traducción literal. ambas no tienen diferencias en el significado.

12) “十分重要”: Se corresponde a “es importante que” en el texto original. En TA se traduce como “重要的是”. En comparación con la TA, la TH destaca el nivel de importancia, con significado “muy importante”, que no se interpreta en el texto original. Entonces, aquí la TA es más adecuadas.

En resumen, entre las 12 diferencias de traducción humana y traducción automática,

a. La TA es mejor: (10), (12)

b. La TH es mejor: (3), (5);

c. La TA y la TH ambas son adecuadas:(1), (2), (4), (6), (7), (8), (9), (11);

d. La TA ni la TH son adecuada: 0

Entonces, entre 12 diferencias, el número de los casos de que la TA es mejor es 2; el de los casos de que la TH es mejor es 2; el de los casos de que ambas son adecuadas es 8.

**La quinta parte:**

En la quinta parte, el contenido principal es la introducción de dos tipos de pruebas para detectar la infección de la TB. Una es el análisis de sangre, que se puede realizar en una visita. La otra es prueba cutánea y necesita volver al médico para revisar su reacción.

## Figura 25

*Comparación de la traducción humana y automática de la quinta parte*

Sentence 5	BLEU	Length ratio	Text
Human	100.00	1.00	结核病感染检测有两种类型。您的医疗服务提供者可能会推荐结核血检，因为它可以在一次就诊中完成，而且如果您过去曾接种过结核病疫苗，那么它更准确。血液检测测量您的免疫系统对引起结核病的病菌的反应。在进行皮试时，要用一根小针将一些检测材料放在皮肤下。您需要在2-3天内回到您的医疗保健医生那里，看看是否有反应。
Machine	35.69	0.96	有两种检测结核病感染的方法。您的医疗保健提供者可能会建议分析结核病检测血液，因为它可以在一次访问中完成如果您过去接种过结核病疫苗，那就更准确了。血液测试测量您的免疫系统对引起结核病的微生物的反应。在结核菌素皮试中，用小针？在皮肤下放一些分析材料。您必须在2或3天内返回您的医疗保健提供者，看看是否有反应。

Según la figura 25, la tasa de la similitud de traducción humana y traducción automática es de 35.69%. De acuerdo con las palabras marcadas en rojo, se comparan y se analizan las diferencias entre las dos versiones traducidas:

- 1) “类型”：Se corresponde a “dos tipos” en el texto original. En TA se traduce como “方法”. Las dos son sinónimos y no tienen diferencias.
- 2) “服务”：Se corresponde a “atención” en el texto original. En TA se traduce como “保健”. Las dos son sinónimos y, además, en el texto posterior de TH se utiliza esta palabra.
- 3) “推荐”：Se corresponde a “podría recomendar” en el texto original. En TA se traduce como “建议”. Las dos traducciones son adecuadas.
- 4) “血检”：Se corresponde a “el análisis de sangre” en el texto original. En TA se traduce como “分析检测血液”, que es una traducción muy literal. La TH es mejor, es que es más profesional y concisa.
- 5) “就诊”：Se corresponde a “una visita” en el texto original. En TA se traduce como “访问”, que es una traducción muy literal. Aquí la TA no es muy adecuada, es que se trata de una visita con el médico, como una consulta. Si se traduce de esta manera, no se refleja el contexto sanitario.
- 6) “那么”：No hay contenido para corresponder. Es una componente para la estructura de frase de condición. En TA se aplica “那就” para formar la relación. Las dos no tienen diferencias.
- 7) “病菌”：Se corresponde a “los microbios” en el texto original. En TA se traduce como “微生物”, que no es exacta, es que los microbios en la ciencia médica aquí se refieren a los gérmenes o virus, mientras que la TA es una palabra en el ámbito biológico.
- 8) “一根”：Se corresponde a “una aguja” en el texto original. En TA no se traduce el clasificador, así no es precisa.
- 9) “检测材料”：Se corresponde a “material de análisis” en el texto original. En TA se traduce como “分析材料”. Las dos locuciones son sinónimos.
- 10) “需要”：Se corresponde a “deberá” en el texto original. En TA se traduce como “必须”. La diferencia entre las dos radica en que la TA destaca la obligatoriedad y la

necesidad de volver al médico. Las dos son comprensibles.

11) “2-3 天”: Se corresponde a “2 o 3 días” en el texto original. En TA se traduce como “2 或 3 天”, que es una traducción muy literal. Pero, las dos no tienen diferencia en el significado.

12) “回到”: Se corresponde a “volver a” en el texto original. En TA se traduce como “返回”. Son sinónimos.

13) “那里”: No hay contenido para corresponder. Es un componente para combinar con la palabra “volver”, destacando el sitio.

En resumen, entre las 13 diferencias de traducción humana y traducción automática,

a. La TA es mejor: 0

b. La TH es mejor: (4), (5), (7), (8);

c. La TA y la TH ambas son adecuadas:(1), (2), (3), (6), (9), (10), (11), (12), (13);

d. La TA ni la TH son adecuada: 0

Entonces, entre 13 diferencias, el número de los casos de que la TA es mejor es 0; el de los casos de que la TH es mejor es 4; el de los casos de que ambas son adecuadas es 9.

### La sexta parte:

En la sexta parte, el contenido principal es el análisis del resultado de la prueba de la piel. En caso de presentar el resultado positivo, usted está infectado por los gérmenes de la tuberculosis, pero hace falta hacer otras pruebas, como una radiografía de tórax o análisis de flema para comprobar si se desarrollan a la enfermedad.

### Figura 26

*Comparación de la traducción humana y automática de la sexta parte*

Sentence 6	BLEU	Length ratio	Text
Human	100.00	1.00	皮试结果若呈阳性会怎样? 阳性皮试结果通常表示您已经受到受结核病菌感染, 但未必表示您已经患有活动期结核病。您需接受其它测试, 例如胸部X光透视或痰(痰液)采样等, 藉以查明您是否患有活动期结核病。
Machine	8.16	1.02	如果皮肤测试结果呈阳性怎么办? 皮肤测试的阳性结果通常意味着您感染了结核病病菌, 但这并不完全意味着您患有结核病。需要其他检查, 如胸部X光或痰(痰)样本, 以确定您是否患有结核病。

Según la figura 26, la tasa de la similitud de traducción humana y traducción automática es de 8.16%. De acuerdo con las palabras marcadas en rojo, se comparan y se analizan las diferencias entre las dos versiones traducidas:

1) “皮试”: Se corresponde a “la prueba de la piel” en el texto original. En TA se traduce como “皮肤测试”, que es el nombre completo de la prueba de la piel. Las dos no se diferencian en el significado.

2) “会怎样”: Se corresponde a “Qué pasa” en el texto original. En TA se traduce como “怎么办”. Las dos traducciones interpretan la misma pregunta.

3) “表示”: Se corresponde a “significa” en el texto original. En TA se traduce como “意



味着”。 Las dos son sinónimos.

4) “受到感染”: Se corresponde a “está infectado” en el texto original. En TA se traduce como “感染了”. La TH es de voz pasiva, mientras que la TA es una expresión coloquial, que no satisface los principios de gramática. En una instrucción para pacientes, es mejor la TH.

5) “未必”: Se corresponde a “no significa precisamente” en el texto original. En TA se traduce como “并不完全”, que es muy literal. La TH enfatiza la posibilidad de tener la enfermedad, mientras que la TA destaca el nivel. Entonces, la traducción automática de esta frase no es muy precisa.

6) “接受测试”: Se corresponde a “Se necesitan otras pruebas ”en el texto original. En TA se traduce como “需要其他检查”, donde existen dos errores: por un lado, falta el sujeto de usted en esta frase, y por otro lado, no se puede usar “其他”. En chino, “其他” se refiere a otras personas, mientras que “其它” significa otras cosas. Entonces, se debe usar “其它”.

7) “例如”: Se corresponde a “como” en el texto original. En TA se traduce como “如” para poner ejemplos. Las dos son sinónimos y no tienen ninguna diferencia.

8) “采样”: Se corresponde a “una muestra” en el texto original. En TA se traduce como “样本”. Las dos ambas son adecuadas.

9) “藉以”: Se corresponde a “para” en el texto original. En TA se traduce como “以”. Aquí la TA es más adecuada, es que la palabra “藉以” se usa con frecuencia en la obra tradicional, que es un poquito redundante.

10) “查明”: Se corresponde a “averiguar” en el texto original. En TA se traduce como “确定”. Aquí la TA es más precisa, es que “查明” se usa más para identificar un caso o averiguar una verdad. En chino es más natural “查明”.

11) “活动期结核病”: Se corresponde a “la enfermedad de la TB” en el texto original. En TA se traduce como “结核病”. La TH es más profesional, que destaca que, en caso de los gérmenes activos, el paciente se enferma. Pero, las dos no tienen diferencia en el significado.

En resumen, entre las 11 diferencias de traducción humana y traducción automática,

a. La TA es mejor: (9), (10);

b. La TH es mejor: (4), (5), (6);

c. La TA y la TH ambas son adecuadas:(1), (2), (3), (7), (8), (10);

d. La TA ni la TH son adecuada: 0

Entonces, entre 11 diferencias, el número de los casos de que la TA es mejor es 2; el de los casos de que la TH es mejor es 3; el de los casos de que ambas son adecuadas es 6.

## **La séptima parte**

En la séptima parte, el contenido principal consiste en los casos que usted está tiene la necesidad de hacer la prueba de la piel para comprobar si está enfermo/a, como en caso de haber viajado por las zonas menos avanzadas, en caso de inyectarse drogas ilegales, en caso de tener sistema inmunológico débil, o en caso de tener síntomas parecidos.

**Figura 27***Comparación de la traducción humana y automática de la séptima parte*

Sentence 7	BLEU	Length ratio	Text
Human	100.00	1.00	应该于何时接受皮试？若存在以下情况，您应接受测试：？存在活动期结核病症状？曾经与活动期结核病患者有长时间接触？患有会削弱人体免疫系统的病症，例如艾滋病毒感染（HIV）或糖尿病等。？曾经注射过非法药品？最近从某个结核病发病率高的地区来到美国，这些地区包括拉丁美洲、加勒比海地区、非洲、亚洲或东欧地区。？曾在护养院、无家可归者救济所、监狱或其它集体环境中工作或住宿。
Machine	11.42	0.99	我应该什么时候做皮肤测试？如果出现以下情况，应进行结核病检测：？他有结核病的症状。？他花了很多时间和一个患有结核病的人在一起。？有条件吗？削弱您的免疫系统，如艾滋病毒或糖尿病？非法药物已被注射？他刚从拉丁美洲、加勒比、非洲、亚洲或东欧等结核病高发地区来到美国。？他曾在养老院、无家可归者收容所、监狱或其他人群中心工作或度过一段时间。

Según la figura 27, la tasa de la similitud de traducción humana y traducción automática es de 11.42%. De acuerdo con las palabras marcadas en rojo, se comparan y se analizan las diferencias entre las dos versiones traducidas:

- 1) “于何时”: Se corresponde a “Cuándo” en el texto original. En TA se traduce como “什么时候”. Las dos interpreta la misma pregunta y no tienen diferencia.
- 2) “存在”: Se corresponde a “si” en el texto original. En TA se traduce como “出现”. Las dos traducciones expresan “si aparecen los casos siguientes”, y son de mismo significado.
- 3) “存在”: Se corresponde a “tiene síntomas” en el texto original. En TA se traduce como “有”. La TH no es perfecta, es que el sujeto de “存在”, desde la gramática, es síntomas, pero en el texto original, el sujeto es usted. La TA también tiene problema, porque el sujeto de “有” es la tercera persona, pero en la traducción posterior, se utiliza la segunda persona usted. Entonces, la traducción buena consiste en “您有”.
- 4) “患者”: Se corresponde a “alguien” en el texto original. En TA se traduce como “人”, que es más general y coloquial. Entonces, la TH es más exacta.
- 5) “有长时间接触”: Se corresponde a “Ha pasado mucho tiempo” en el texto original. En TA se traduce como “花很多时间再一起”, que es muy literal. Así, la TH es mejor.
- 6) “患有病症”: Se corresponde a “tiene una condición” en el texto original. En TA se traduce como “有条件吗”, que es una traducción literal, pero es falsa. La TH es correcta.
- 7) “曾经”: Se corresponde a “se ha inyectado” en el texto original. En TA se traduce como “已经”. La TH destaca tener la experiencia de inyección de drogas, mientras que la TA enfatiza que la acción de inyección ya termina. Entonces, la TH es más adecuada.
- 8) “药品”: Se corresponde a “drogas” en el texto original. En TA se traduce como “药物”. Las dos palabras son sinónimos.
- 9) “发病率高的”: Las dos traducciones contienen mismo significado. La TA es “高发”.
- 10) “最近”: Se corresponde al tiempo de “Ha venido” en el texto original. En TA se traduce como “刚”. El plazo de TH es más largo que TA. Desde mi punto de vista, la TH es más adecuada aquí.
- 11) “加勒比海”: Se corresponde a “Caribe” en el texto original. En TA se traduce como “加勒比”. Aunque la TH destaca el mar, las dos traducciones son correctas.
- 12) “护养院”: Se corresponde a “un hogar de ancianos” en el texto original. En TA se traduce como “养老院”, que se usa más en China. Entonces, es mejor aplicar la TA.
- 13) “救济所”: Se corresponde a “un refugio” en el texto original. En TA se traduce como

“收容所”. Las dos son sinónimas.

14) “其它集体环境”: Se corresponde a “otro centro para grupos de personas” en el texto original. En TA se traduce como “其他人群中心”, que significa “otro centro de personas”, además, donde se usa falsamente la palabra “其他”. Entonces, aquí la TH es más adecuada.

En resumen, entre las 14 diferencias de traducción humana y traducción automática,

a. La TA es mejor: (12);

b. La TH es mejor: (4), (5), (6), (7), (10), (14);

c. La TA y la TH ambas son adecuadas:(1), (2), (8), (11), (13), (9);

d. La TA ni la TH son adecuada: (3)

Entonces, entre 14 diferencias, el número de los casos de que la TA es mejor es 1; el de los casos de que la TH es mejor es 6; el de los casos de que ambas es adecuadas son 6; el número de que ambas son impropias es 1.

## La octava parte

En la octava parte, el contenido principal es los posibles efectos secundarios de los medicamentos para tratar la TB. Es cierto que los efectos secundarios no son comunes, pero los pacientes tienen la posibilidad de tener pérdida del apetito, orina oscura, vómito, náuseas, etc. después de tomar las medicinas.

## Figura 28

*Comparación de la traducción humana y automática de la octava parte*

Sentence 8	BLEU	Length ratio	Text
Human	100.00	1.00	潜伏结核感染(LTBI)治疗药物可能会产生哪些副作用? 此类药物的副作用十分罕见, 但可能包括: ? 恶心? 呕吐? 尿液呈深色? 皮疹三? 皮肤和 (或) 眼睛呈黄色? 食欲不振? 双手和 (或) 双足麻木
Machine	6.59	0.81	结核病药物可能有哪些副作用? 这些药片的副作用并不常见, 但它们可能包括: ? 恶心。? 呕吐。? 深色尿液? 皮肤上的皮疹? 皮肤或眼睛发黄? 食欲不振? 手或脚麻木

Según la figura 28, la tasa de la similitud de traducción humana y traducción automática es de 6.59%. De acuerdo con las palabras marcadas en rojo, se comparan y se analizan las diferencias entre las dos versiones traducidas:

1) “会产生”: Se corresponde a “algunos posibles efectos” en el texto original. En TA se traduce como “可能有”. Las dos describe una posibilidad de aparecer los efectos secundarios y son de mismo significado.

2) “此类药物”: Se corresponde a “con estas pastillas” en el texto original. En TA se traduce como “这些药片”. Es mejor la TH, porque la TA es más concreta, pero la verdad es que no sabe cuáles son las pastillas concretamente.

3) “十分罕见”: Se corresponde a “no son comunes” en el texto original. En TA se traduce como “并不常见”. Es mejor la TA, es que la TH destaca la rareza de los efectos secundarios, pero en el texto original solo describe que los efectos no son comunes.

4) “皮疹三”: Se corresponde a “sarpullido” en el texto original. En la TA se traduce como

“皮肤上的皮疹”, que es redundante. La TH tampoco es falsa, y la mejor traducción es “皮疹”.

5) “呈黄色”: Se corresponde a “amarillentos” en el texto original. En la TA se traduce como “发黄”. Las dos son sinónimas.

6) “双手”: Se corresponde a “manos” en el texto original. En la TA se traduce como “手”. La TH destaca dos manos, que es más rigurosa.

7) “双足”: Se corresponde a “pies” en el texto original. En la TA se traduce como “脚”. La razón es como el sexto punto, así la TH es más precisa.

En resumen, entre las 7 diferencias de traducción humana y traducción automática,

- a. La TA es mejor: (3);
- b. La TH es mejor: (2), (6), (7);
- c. La TA y la TH ambas son adecuadas:(1), (5);
- d. La TA ni la TH son adecuada: (4);

Entonces, entre 7 diferencias, el número de los casos de que la TA es mejor es 1; el de los casos de que la TH es mejor es 3; el de los casos de que ambas son adecuadas son 2; el número de que ambas son impropias es 1.

## La novena parte

En la novena parte, el contenido principal es la introducción de las pruebas especiales para diagnosticar la infección a fin de comprobar si está contagiada o no. Hay dos diferentes maneras: el análisis de sangre y una prueba de tuberculina.

### Figura 29

*Comparación de la traducción humana y automática de la novena parte*

Sentence 9	BLEU	Length ratio	Text
Human	100.00	1.00	我没觉得身体不适, 如何判断被感染呢? 结核感染不会有任何症状, 所以我们会通过特殊检查来诊断。您可以在当地结核服务机构做检查, 也可以得到他们的相关帮助。检查包括: ?通过γ干扰素释放试验血检 (IGRA 或 QuantiFERON TB Gold Plus ?) 可判断您是否感染了结核病。?通过结核菌素皮试 (又称作 TST 或 结核菌素试验), 在前臂进行量注射。注射 2 至 3 天之后再次检查, 根据身体反应 (红肿) 判断是否被感染结核病。?通过胸部 X 光可判断肺部是否有结核。有时 X 光还可查出之前的结核, 即使您并不知道之前患过结核病。
Machine	15.39	0.92	我不觉得恶心, ?我怎么知道我不是被传染了? 结核病感染不会引起任何症状, 因此我们必须进行特殊测试来诊断结核病感染。可以通过当地的结核病服务机构进行或协调测试。这些测试包括: ?一项名为γ干扰素释放试验 (IGRA 或 QuantiFERON TB Gold Plus ?) 的血液检查可以证明您是否感染了结核病。?结核菌素皮肤试验 (也称为 TST 或 Mantoux 试验) 是一种小注射, 正好放置在前臂内侧的皮肤下。您必须在测试后 2 至 3 天内返回, 以便能够评估任何表明结核病感染的反应 (肿胀和发红)。

Según la figura 29, la tasa de la similitud de traducción humana y traducción automática es de 15.39%. De acuerdo con las palabras marcadas en rojo, se comparan y se analizan las diferencias entre las dos versiones traducidas:

- 1) “没觉得身体不适”: Se corresponde a “No me siento enfermo” en el texto original. En TA se traduce como “不觉得恶心”, que no es correcta, es que la traducción automática significa tener náuseas, mientras que el texto original se refiere a que está mal.
- 2) “如何判断”: Se corresponde a “cómo puedo saber si” en el texto original. En TA se traduce como “怎么知道是不是”, que es una traducción literal, pero son de mismo

significado.

3) “有”: Se corresponde a “no causa ningún síntoma” en el texto original. En TA se traduce como “引起”. Las dos en este contexto son adecuadas.

4) “所以”: Se corresponde a “por lo cual” en el texto original. En TA se traduce como “因此”. Las dos traducciones son sinónimas.

5) “会”: Se corresponde a “tenemos que” en el texto original. En TA se traduce como “必须”. En comparación con TH, la TA destaca la necesidad de realizar pruebas. Entonces, la TA puede interpretar el tono y la urgencia, así es más adecuada.

6) “做检查”: Se corresponde a “Las pruebas se pueden hacer o coordinar” en el texto original. En TA se traduce como “协调测试”, que es literal, pero a esta frase traducida por TA falta el objeto. Entonces, prefiero la TH.

7) “试验血检”: Se corresponde a “Un examen de sangra” en el texto original. En TA se traduce como “血液检查”, que es el nombre completo de esta prueba. Las dos traducciones se pueden comprender.

8) “可判断”: Se corresponde a “puede demostrar” en el texto original. En TA se traduce como “可以证明”. La diferencia entre las dos radica en que la TH se refiere a que puede juzgar si se ha contagiado con B, mientras que la TA se refiere a que puede comprobar si se ha contagiado. Pero las dos son comprensibles.

9) “皮试”: Se corresponde a “Una prueba cutánea” en el texto original. En TA se traduce como “皮肤试验”, que es el nombre completo de la prueba. Las dos traducciones son iguales en el significado.

10) “又称作”: Se corresponde a “conocida también como” en el texto original. En TA se traduce como “也称为”. Las dos son sinónimas.

11) “结核菌素试验”: Se corresponde a “prueba Mantoux” en el texto original. En TA se traduce como “Mantoux 试验”, que mantiene el nombre de inglés. Debería traducir el nombre extranjero de esta prueba, así la TH es mejor.

12) “量注射”: Se corresponde a “una pequeña inyección” en el texto original. En TA se traduce como “小注射”. Las dos ninguna son adecuadas. En TH falta la traducción de la palabra “pequeña”, mientras que en TA no se usa “小” para describir una inyección, el cual es una traducción literal. Entonces, la traducción adecuada es “少量注射”.

13) “在前臂进行”: Se corresponde a “colocada justo bajo la piel de la parte interna del antebrazo” en el texto original. En TA se traduce como “正好放置在前臂内侧”. Las dos traducciones no son perfectas. Por un lado, en TH falta la traducción de “la parte interna”. Por otro lado, la TA es muy literal, la cual se puede entender, pero no es perfecta. Entonces, la traducción precisa es “在前臂内侧进行”.

14) “身体的反应”: Se corresponde a “cualquier reacción” en el texto original. En TA se traduce como “反应”. Las dos no tienen diferencias en el significado.

15) “红肿”: Se corresponde a “hinchazón y enrojecimiento” en el texto original. En TA se traduce como “肿胀和发红”, que es descripción completa. La TH es una abreviatura, pero las dos son comprensibles.

16) “判断”: Se corresponde a “evaluar” en el texto original. En TA se traduce como “评估”, que es muy literal. La TA solo se refiere a evaluar reacción, mientras que la TH destaca el objetivo de evaluar reacción, que es evaluar si se contagia la TB. Así, prefiero la TH.

17) “返回”: Se corresponde a “deberá regresar” en el texto original. En TA se traduce

como “再次”. Las dos no tienen diferencias en el significado.

18) “得到他们的相关帮助”: No hay contenido para corresponder en el texto original, es una parte añadida en la TH. No es muy adecuada.

19) “通过胸部 X 光可判断肺部是否有结核。有时 X 光还可以查出之前的结核，即时您并不知道之前患过结核病。”: No hay contenido para corresponder en el texto original. Es una versión oficial en el sitio web, así puede ser parte china agregada.

En resumen, entre las 19 diferencias de traducción humana y traducción automática,

a. La TA es mejor: (1), (18), (19);

b. La TH es mejor: (5), (6), (11), (16);

c. La TA y la TH ambas son adecuadas: (2), (3), (4), (7), (8), (9), (10), (14), (15), (17);

d. La TA ni la TH son adecuadas: (12), (13);

Entonces, entre 19 diferencias, el número de los casos de que la TA es mejor es 3; el de los casos de que la TH es mejor es 4; el de los casos de que ambas son adecuadas son 10; el número de que ambas son impropias es 2.

## La décima parte

En la décima parte, el contenido principal es la introducción de la vacuna de Bacilo Calmette-Guérin (BCG). Se trata de una vacuna de bacilo vivo para la prevención en algún sentido. Aunque no se puede prevenir todo el riesgo de infectarse, se puede eliminar la posibilidad que la infección sea leve. Tarda seis y doce semanas reproducir el efecto y la eficacia durará de cinco a quince años.

### Figura 30

*Comparación de la traducción humana y automática de la décima parte*

Sentence 10	BLEU	Length ratio	Text
Human	100.00	1.00	卡介苗疫苗 (BCG) 是什么? 卡介苗疫苗是一种活性疫苗, 可有限预防结核病。若接触到结核病病菌, 卡介苗疫苗不能保护您免受感染, 但却极有可能把感染变成轻微的局部感染, 而不是威胁生命的重病。卡介苗疫苗需要 6 至 12 周时间来产生免疫力保护, 预防结核病的免疫能力可能只有 50 - 60%, 有些人体内的疫苗会在 5 - 15 年间失去效力。
Machine	18.48	1.21	卡介苗是什么? 卡介苗是一种活杆菌疫苗, 可提供一定程度的结核病保护。卡介苗并不能防止一个人在接触结核病细菌 (病菌) 时被感染, 但它会大大增加一个人感染轻微和局部感染的可能性, 而不是一种压倒性的、可能致命的疾病。卡介苗可能需要 6 至 12 周才能产生免疫 (保护) 作用。卡介苗疫苗只能提供 50 - 60% 的结核病免疫力, 在某些人中, 疫苗会随着时间的推移而失效, 有时会在 5 - 15 天内失效。

Según la figura 30, la tasa de la similitud de traducción humana y traducción automática es de 18.48%. De acuerdo con las palabras marcadas en rojo, se comparan y se analizan las diferencias entre las dos versiones traducidas:

1) “活性”: Se corresponde a “de bacilo vivo” en el texto traducido. En TA se traduce como “活杆菌”, que es una traducción literal. El texto original describe la naturaleza de la vacuna, así la TH es más adecuada.

2) “有限预防”: Se corresponde a “cierto grado de protección contra” en el texto original. En TA se traduce como “提供一定程度的保护”. La TH es más concisa y de términos médicos.

- 3) “保护您免受感染”: Se corresponde a “evita que la persona resulte infectada” en el texto original. En TA se traduce como “防止一个人被感染”. Las dos tienen mismo significado.
- 4) “极有可能把感染变成”: Se corresponde a “aumenta en mucho la probabilidad” en el texto original. En TA se traduce como “大大增加一个人感染可能性”, que es muy literal. En cuanto a la estructura y el tono natural, la TH es mucho mejor.
- 5) “威胁生命的”: Se corresponde a “brumadora” en el texto original. En TA se traduce como “压倒性的”, que es el significado original de la palabra. Pero describiendo una enfermedad, es mejor la TH, más natural y exacta.
- 6) “重病”: Se corresponde a “posiblemente fatal” en el texto original. En TA se traduce como “致命的疾病”. Las dos son de mismo significado, pero la TH es más concisa y breve.
- 7) “预防能力”: Se corresponde a “brindar la inmunidad” en el texto original. En TA se traduce como “提供免疫力”, que es una traducción literal. Prefiero la TH.
- 8) “有些人体内”: Se corresponde a “en algunas personas” en el texto original. En TA se traduce como “在某些人中”. La TH destaca que la inmunidad está en el cuerpo de algunas personas, mientras que la TA es literal sin incluir el significado que la frase original quería expresar.
- 9) “年间”: Se corresponde a “años” en el texto original. En TA se traduce como “天内”, que es falsa completamente, es que es la traducción de “días”.
- 10) “失去效力”: Se corresponde a “pierde su eficacia” en el texto original. En TA se traduce como “失效”. Las dos son de mismo significado.

En resumen, entre las 10 diferencias de traducción humana y traducción automática,

- a. La TA es mejor: 0
- b. La TH es mejor: (1), (2), (4), (5), (6), (7), (8), (9);
- c. La TA y la TH ambas son adecuadas:(3), (10);
- d. La TA ni la TH son adecuada:0

Entonces, entre 10 diferencias, el número de los casos de que la TA es mejor es 0; el de los casos de que la TH es mejor es 8; el de los casos de que ambas son adecuadas son 2; el número de que ambas son impropias es 0.

### **La undécima parte:**

En la última parte, se proponen unos asuntos a que deben prestar atención sobre la BCG y otras vacunas. Por una parte, es necesario recibir todas las vacunas en un día o esperar por lo menos cuatro semanas después de inyección de la vacuna BCG para recibir otras vacunas. También hay otras vacunas como vacunas contra fiebre amarillas, paperas, rubéola, sarampión, la tifoidea, etc.

### **Figura 31**

*Comparación de la traducción humana y automática de la undécima parte*

Sentence 11	BLEU	Length ratio	Text
Human	100.00	1.00	卡介苗疫苗和其他活性疫苗。前往海外旅游而需要接种超过一种活性疫苗的人士，应该：a)在同一天接种所有活性疫苗，或b)接种卡介苗疫苗和其他活性疫苗之间等候最少四星期。这样做是让身体的免疫(防卫)系统最佳的机会，提供所需的免疫力(保护)。其他的活性疫苗包括：口服伤寒疫苗，黄热病疫苗，腮腺炎、麻疹、德国麻疹混合疫苗。卡介苗疫苗和口服小儿麻痹疫苗可在四星期内接种。
Machine	11.86	1.09	卡介苗和其他活杆菌/病毒疫苗。出国旅行并在旅行前需要一种以上活杆菌/病毒疫苗的人，应在同一天接种所有活杆菌/病毒疫苗，或(b)在卡介苗和其他活杆菌/病毒疫苗之间至少等待4周。这为身体的免疫(防御)系统提供了更大的机会来提供必要的免疫(保护)。其他细菌/病毒疫苗包括：口服伤寒、黄热病和腮腺炎、麻疹和风疹(MMR)疫苗。卡介苗和口服抗脊髓灰质炎药物可在4周内服用。

Según la figura 31, la tasa de la similitud de traducción humana y traducción automática es de 11.84%. De acuerdo con las palabras marcadas en rojo, se comparan y se analizan las diferencias entre las dos versiones traducidas:

- 1) “前往海外旅行”: Se corresponde a “viajen al extranjero” en el texto original. En TA se traduce como “出国旅行”, que es más concisa y breve. Prefiero la traducción automática.
- 2) “超过”: Se corresponde a “más de” en el texto original. En TA se traduce como “以上”. Las dos son sinónimos y ni tienen diferencia.
- 3) “人士”: Se corresponde a “Las personas” en el texto original. En TA se traduce como “人”. La TH es más formal, pero las dos son comprensibles.
- 4) “应该”: Se corresponde a “debería” en el texto original. En TA se traduce como “应”. Las dos son sinónimos. En chino se omiten palabras con frecuencia sin influenciar el significado de la frase.
- 5) “等候最少四星期”: Se corresponde a “esperar por lo menos 4 semanas” en el texto original. En TA se traduce como “至少等待 4 周”, que es más formal y concisa a pesar de tener el mismo significado la TA y la TH.
- 6) “这样做是让”: Se corresponde a “Con ello” en el texto original. En TA se traduce como “这为”, que es más breve.
- 7) “最佳”: Se corresponde a “la mayor oportunidad” en el texto original. En TA se traduce como “更大的机会”, que es muy literal. Prefiero la TH, es que cambia la forma para ajustar al hábito chino y es más natural
- 8) “所需的”: Se corresponde a “necesaria” en el texto original. En TA se traduce como “必要的”. Las dos son sinónimos.
- 9) “活性”: Se corresponde a “vacuna de bacilo/virusivo” en el texto original. En TA se traduce como “细菌/病毒”, que es una falsa traducción.
- 10) “德国麻疹”: Se corresponde a “sarampión” en el texto original. En TA se traduce como “麻疹”. Las dos se refiere a una misma enfermedad.
- 11) “混合疫苗”: Se corresponde a “rubéola” en el texto original. En TA se traduce como “风疹疫苗”. La TH no es muy exacta, así prefiero la TA.
- 12) “小儿麻痹疫苗”: Se corresponde a “la antipoliomielítica” en el tetxo original. En TA se traduce como “抗脊髓灰质炎药物”, “小儿麻痹” y “抗脊髓灰质” ambas se refiere a “antipoliomielítica”, pero la TA se refiere a una medicina sino una vacuna. Así, la TH es más precisa.
- 13) “接种”: Se corresponde a “se pueden administrar” en el texto original. En TA se traduce como “4 周内服用”, malentendiendo la vacuna por la medicina.

En resumen, entre las 13 diferencias de traducción humana y traducción automática,



- a. La TA es mejor: (1), (5), (6), (11);
- b. La TH es mejor: (7), (8), (12), (13);
- c. La TA y la TH ambas son adecuadas:(2), (3), (4), (8), (10);
- d. La TA ni la TH son adecuada:0

Entonces, entre 13 diferencias, el número de los casos de que la TA es mejor es 4; el de los casos de que la TH es mejor es 4; el de los casos de que ambas son adecuadas son 5; el número de que ambas son impropias es 0.

Después de analizar cada parte de la traducción de manera detenida, se establece una tabala para enumerar la calidad de cada parte de las disimilitudes. Como ya se encuentra de primero que el resultado de similitud de la traducción humana y la traducción automática es de 17.77, se excluyen las zonas repetitivas marcadas en rojo que se refiere a la diferencia, por lo que en base al análisis anterior, entre las dos traducciones hay 146 disimilitudes, entre las que la cantidad de que la calidad de la traducción automática es de 23, la de que la calidad de la traducción humana es de 51, la de que ambas son adecuadas es de 68, y la de ninguna es adecuada es de 4. Se puede ver claramente los resultados de cada parte en la figura 32.

### Figura 32

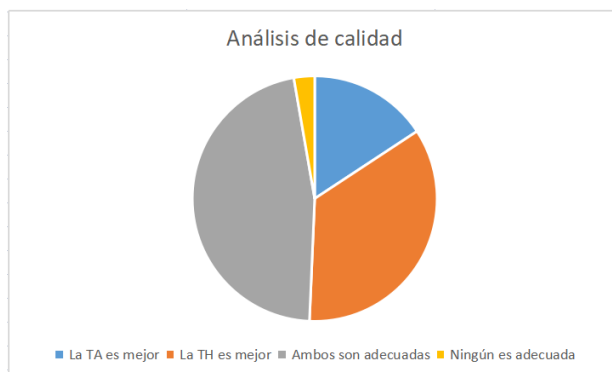
#### *Resumen de la calidad de las disimilitudes*

Diferentes partes	Disimilitudes	La TA es mejor	La TH es mejor	Ambos son adecuadas	Ningún es adecuada
La primera parte	24	7	7	10	0
La segunda parte	13	1	5	7	0
La tercera parte	10	1	5	4	0
La cuarta parte	12	2	2	8	0
La quinta parte	13	0	4	9	0
La sexta parte	11	2	3	6	0
La séptima parte	14	1	6	5	1
La octava parte	7	1	3	2	1
La novena	19	3	4	10	2
La décima parte	10	1	8	2	0
La undécima parte	13	4	4	5	0
Total	146	23	51	68	4

Por medio de la evaluación automática, se ve que el resultado de similitud no puede representar la calidad de la traducción automática. Entre 146 diferencias, 91 (51+68) diferencias no significan que la traducción automática es falsa o impropia, ocupando el 62.33% de la totalidad, mientras que solo entre el 37.67%, la traducción automática tiene problemas. En la figura 34, se puede comparar el porcentaje de la calidad de las disimilitudes entre la traducción humana y la traducción automática.

### Figura 33

#### *Porcentaje de la calidad de las disimilitudes*



### 3.3.2 Evaluación humana

Para realizar la evaluación humana de la traducción automática, se aplica la métrica TAUS (Translation Automation User Society), que se ha introducido de detalle en el segundo capítulo. Se analiza cada parte del corpus para localizar los errores, y se clasifica los errores por cinco categorías: grave, importante, poco importante, neutro, excelente (Sánchez, 2020). Los principios para la asignación de valores a la tipología de errores están basados en la tabla 9.

Antes que nada, aunque ya se explican en el segundo capítulo, cabe mencionar otra vez las siglas aplicadas en los siguientes análisis:

T<sub>EG</sub> es Total de Errores Graves;

T<sub>EI</sub> es Total de Errores Importantes;

T<sub>EPI</sub> es Total de Errores Poco Importantes;

T<sub>EN</sub> es Total de Errores Neutros;

T<sub>ME</sub> es Total de Menciones Excelentes;

VP<sub>EG</sub> es Valor de Penalización de Errores Graves;

VP<sub>EI</sub> es Valor de Penalización de Errores Importante;

VP<sub>EPI</sub> es Valor de Penalización de Errores Poco Importantes;

V<sub>ME</sub> es Valores de las Menciones Excelentes;

TPC es Valor de Penalización por Categoría;

TPO es Total de Palabras del Texto Original;

PPP es el Porcentaje de Penalización por Palabra. (Sánchez, 2020)

#### En la primera parte:

Errores importantes:

1) “LTBI”: En la traducción automática no se traduce LTBI, que el nombre completo es latente TB infección, es decir infección de latente tuberculosis (潜伏期结核病), lo que puede confundir y obstaculizar al usuario, es que no sabe qué significa sin traducción china.

2) “它”: La frase original es “puede contagiar a otras personas con los gérmenes”. El sujeto de esta frase es usted o una persona que tiene la enfermedad de la TB, pero aquí se equivoca y se traduce el sujeto como el germen o virus, que es completamente falsa.

Errores poco importantes:

1) “如果不治疗”: La frase original es “si no se trata”, y la frase completa debe ser “si no se trata la enfermedad”, así el objetivo de tratar es la enfermedad. Pero en la traducción automática falta el sujeto y objeto, y esta frase no es completa desde el punto de vista gramática.

### **En la segunda parte:**

Errores poco importantes:

1) “结核病病菌在空气中传播，患有活动性结核病的人咳嗽，打喷嚏，大笑，说话，或者唱歌”: La frase original es “Los gérmenes se propagan cuando una persona.”. La relación lógica condición es falsa, mientras que la traducción correcta es “当患者咳嗽...，病菌传播”. Los usuarios pueden entender, pero de verdad hay un error sintáctico.

Errores Neutros:

“人”: Se corresponde a la palabra “personas” en el texto original. No influencia el significado, pero el estilo no se adapta al ámbito médico.

### **En la tercera parte:**

Errores graves:

1) “也可能发生几次。你们等一下”: La frase original es “esto puede suceder varios años después”. La traducción automática de esta frase es completamente falsa, lo que puede tener implicaciones en la salud.

Errores importantes:

1) “LTBI”: Se explica en el primer punto de la primera parte.  
2) “结合每周服用 12 周的药物”: Se corresponde a “que combina dos medicinas que se toman semanalmente durante 12 semanas”. Esta traducción no es exacta, la traducción concreta es “每周服用 2 种药物，服用 12 周”. Sin expresar el significado en chino del texto original, la traducción puede confundir el uso correcto de la medicina.

Errores poco importantes:

1) “较短的方案”: Se corresponde a “un régimen más corto” en el texto original. Es una traducción literal sin tener en cuenta el contexto. Es mejor “治疗疗程较短的方案”, añadiendo la parte omitida de “un régimen de tratamiento más corto”. En español se puede omitir la parte repetida, pero en chino si no se añade, le dificulta al paciente para comprenderlo.

### **En la cuarta parte**

Errores poco importantes:

1) “存活了很久”: Se corresponde a “permanecen vivos por un largo tiempo” en el texto

original. El significado de la traducción no tiene problema, pero el tiempo del verbo “permanecer vivo” se equivoca. La frase original significa que los gérmenes todavía van a permanecer vivos por un período después de comenzar a tomar la medicina, y no se refiere a que los gérmenes ya han permanecido vivos mucho tiempo.

Errores neutros:

- 1) “死亡”: La palabra española es “mueren”, pero en chino no se usa con frecuencia “死亡” para describir los gérmenes o virus. El uso de “被杀灭” es mejor. Se trata de un problema de estilo sin influenciar el significado.
- 2) “完成您的药物治疗”: Se corresponde a “termine su medicina” en el texto original. Es una traducción literal y es mejor la traducción “服用全部药物”.

### **En la quinta parte:**

Errores poco importantes:

- 1) “分析结核病检测血液”: Se corresponde a “el análisis de sangre de detección de tuberculosis” en el texto original. Es una traducción literal, que dificulta la comprensión de esta prueba. “血检” dos palabras son suficientes.
- 2) “访问”: La palabra española correspondiente es “visita”. En el ámbito médico, “访问” no es muy adecuada, y se usa más en la vida diaria, una visita con profesores, con amigos, etc.
- 3) “微生物”: Se corresponde a la palabra española “microbio”, y la traducción a “微生物” es muy literal. No influye el significado, pero no se ajusta al estilo aquí.
- 4) “小针”: Se corresponde a “una aguja pequeña” en el texto original. La traducción automática omite el cuantificador “una”. “一根小针” es más exacta.

Errores neutros:

- 1) “?”: En la traducción de iFLYTEK aparece este signo, que no existe en el texto original. No cambia el significado original, pero es un problema de forma.

### **En la sexta parte:**

Errores poco importantes:

- 1) “需要其它检查”: Se corresponde a “Se necesitan otras pruebas” en el texto original. La frase original es de voz pasiva, mientras que en la frase traducida no hay un sujeto. Entonces, hace falta añadir el sujeto “您” en la traducción.

### **En la séptima parte:**

Errores importantes:

- 1) “有条件吗? ”: Se corresponde a “tiene una condición” en el texto original. La traducción automática no conecta el contexto, y es de falso significado. La frase original se refiere a que a usted ocurrírsele que debilita el sistema inmunológico. Entonces, es

mejor la traducción “曾出现情况”.

Errores poco importantes:

- 1) “他有”: La frase española es “Tiene síntomas”. No se escribe el sujeto en el texto original, es que la persona de todo el texto es usted, así aquí se omite. Pero el sujeto aplicado en la traducción automática es “él”. Es falsa, y no se corresponde al estilo del texto.
- 2) “他花了很多时间”: La frase española es “Ha pasado mucho tiempo”. La causa es igual como el primer punto.
- 3) “他刚从”: La frase española es “Ha venido a Estados Unidos”. La razón es igual como el primer punto.
- 4) “他曾在”: La frase española es “Ha trabajado o ha pasado tiempo”. La causa es igual que el primer punto.

Errores neutros:

- 1) “非法药物已被注射”: Se corresponde a “Se ha inyectado drogas ilegales” en el texto original. La traducción no tiene problema en el significado ni en la gramática. No se ajusta a la forma de estructura de otras frases. Es mejor mantenerse uniforme con las otras.
- 2) “其他人群中心工作”: Se corresponde a “otro centro para grupos de personas” en el texto original. No influencia el significado original, pero no se utiliza mucho este nombre, y es mejor cambiar por “其它集体环境”.
- 3) “? ”: En el texto original la forma es de “•” para enumerar diferentes casos. La traducción automática no influencia el significado, pero no se corresponde a la forma original.

### **En la octava parte:**

Errores neutros:

- 1) “结核病药物”: Se corresponde a “las medicinas de la enfermedad” en el texto original. Con esta traducción se puede transmitir el significado original, pero se puede ser mejor “治疗结核病的药物” por medio de agregar el contenido omitido en español.

### **En la novena parte:**

Errores importantes:

- 1) “恶心”: Se corresponde a “me siento enfermo” en el texto original. Esta traducción es falsa, y significa que “tener náuseas”. Este error puede inducir malentendidos al paciente, es que esta frase se refiere a que en el caso de que el paciente está mal, tiene necesidad de hacer pruebas.

Errores poco importantes:

- 1) “小注射”: Las palabras españolas correspondientes son “una pequeña inyección”. El adjetivo se usa no para describir el tamaño de la inyección sino para describir la cantidad.

Así, la traducción “小” no es exacta, y es más adecuada “少量注射”.

Errores neutros:

1) “测试”: Se corresponde a “pruebas” en el texto original. Es más exacta usar “检查” en el ámbito médico, porque la “测试” se utiliza en el ámbito académico con más frecuencia.

**En la décima parte:**

Errores poco importantes:

1) “活杆菌”: Se corresponde a “de bacilo vivo” en el texto original. Es una traducción literal. En caso de describir una vacuna, se utiliza “活性的” en el ámbito médico.

2) “压倒性的”: Se corresponde a “brumadora” en el texto original. La traducción automática es muy literal, y solo se traduce el significado del adjetivo. Para describir una enfermedad, se utiliza “重病”, que es más natural.

**En la undécima parte:**

Errores importantes:

1) “卡介苗和口服抗脊髓灰质炎药物”: Se corresponde a “La BCG y la antipoliomielítica” en el texto original. Las dos no son medicinas sino vacunas, por lo que la traducción automática es falsa.

Errores poco importantes:

1) “活杆菌”: Se corresponde a “de bacilo/virus vivo” en el texto original. La razón es igual que el primer punto de la décima parte.

Errores neutros:

1) “提供了更大的机会来提供必要的免疫”: Se corresponde a “brindar la mayor oportunidad de ofrecer la inmunidad” en el texto original. La traducción automática puede transmitir el significado original, pero no es natural en chino. Entonces, es mejor “让身体的免疫 (防卫) 系统处于最佳状态来提供免疫力”.

En resumen, en base a todos los análisis arriba, se cuenta la cantidad de cada categoría de error para hacer cálculos de los valores de penalización: hay uno error grave, siete errores importantes, diecisiete errores poco importantes, diez errores neutros, y cero menciones excelentes. Además, el Total de Palabras del Texto Original (TPC) es de 1082. Como se explica en el segundo capítulo, la penalización de cada error es:

$$VP_{EPI}=0.25$$

$$VP_{EI}=0.5$$

$$VP_{EG}=0.75$$

$$V_{ME}=1$$

Entonces, se hace cálculo de total de penalizaciones por categoría:

$$\begin{aligned} \text{TPC} &= (\text{T}_{\text{EG}} \times \text{VP}_{\text{EPG}}) + (\text{T}_{\text{EI}} \times \text{VP}_{\text{EI}}) + (\text{T}_{\text{EPI}} \times \text{VP}_{\text{EI}}) - (\text{T}_{\text{ME}} \times \text{V}_{\text{ME}}) \\ &= (1 \times 0.75) + (7 \times 0.5) + (17 \times 0.5) - 0 = 12.75 \\ \text{PPP} &= (\text{TPC} \div \text{TPO}) \times 100 = (12.75 \div 1082) \times 100 = 1.178 \text{ (resultado redondeo)} \end{aligned}$$

$$\text{Calidad de la traducción} = 100 - \text{PPP} = 98.882$$

Por último, por medio de la evaluación humana, el resultado de la traducción es 98.882.

De acuerdo con el resultado y el análisis de los errores, se encuentra que en la traducción automática no hay muchos errores en la gramática, en la sintaxis, sino en las expresiones. El problema principal que causa a los errores radica en la traducción literal. El iFLYTEK no tiene capacidad de ajustar las expresiones, la estructura y el uso de palabras según los diferentes hábitos y diversas culturas. Generalmente, el resultado de la calidad de la traducción automática de iFLYTEK es bueno, lo que significa que con el traductor puede satisfacer la necesidad diaria.

#### **4. Análisis y resultados**

En el tercer capítulo, se han analizado los errores y los problemas de cada parte del corpus de la traducción. Se puede encontrar de manera clara la calidad y las características de la traducción automática realizada por iFLYTEK. En este capítulo, se resumen los rasgos principales de iFLYTEK y los comentarios de la métrica BLEU con respecto a la evaluación automática a fin de proponer unos consejos y brindar inspiraciones a los estudios posteriores.

##### **4.1 Características de la TA de iFLYTEK**

El traductor de máquina puede traducir un texto en poco tiempo, es decir, tiene muy alta eficiencia en el rendimiento, por lo que se puede ahorrar mucho tiempo y energía por medio de usarlo. Pero en el proceso de aplicarlo en la realidad aparecen unos problemas, que reducen la calidad del texto traducido. En cuanto al traductor iFLYTEK, después de utilizarlo en la traducción de un corpus en el ámbito sanitario, especialmente en el ámbito de tuberculosis, se encuentran unas desventajas y se explican de manera detallada en este apartado.

###### **4.1.1 Traducción literal**

La traducción literal, según Sara M. Parkinson de Saz (1984), es una técnica de traducción, es decir, una traducción palabra por palabra. Según la definición, se ve que, si se traduce palabra por palabra, es cierto mantener la estructura y el orden de palabras del idioma de origen. Sin embargo, las dos diferentes lenguas consisten en frutos de diferentes culturas y contextos, lo que requiere ajuste y cambio en la traducción. En caso de traducción palabra por palabra, es posible dificultar la comprensión para los usuarios del idioma de destino.

La técnica principal aplicada de iFLYTEK es la traducción literal. Esta traducción automática puede transferir el significado de una frase del idioma origen, mientras que el estilo es muy duro y no es natural. Con la traducción literal también conduce a algunos errores de la relación de oraciones debido a diferentes características del lenguaje español y chino, y puede causar malentendidos para el texto original.

Por un lado, por medio de traducción literal, es fácil hacer la frase china parecer más redundante. Durante el análisis se encuentra que el iFLYTEK obedece el orden de las frases españolas cien por cien para realizar la traducción de las frases chinas. En este caso, se puede entender el significado, pero no es muy natural y necesario. Como, por ejemplo, en cuanto a la frase “La infección significa que ..., pero...”, la traducción china sigue el orden y la estructura “感染意味着..., 但是...”. El lenguaje chino se caracteriza por lo conciso, entonces, es un poquito curioso traducir la definición de la infección como dos frases.

Por otro lado, con traducción literal, se puede equivocar de la relación lógica. En el lenguaje español se utiliza subordinaciones de tiempo o de condición para interpretar la relación lógica, pero las estructuras de las subordinaciones en chino y en español son diferentes. Como, por ejemplo, en la traducción de esta frase “los gérmenes se propagan cuando una persona...”, “结核病传播时, 人咳嗽...”, la relación lógica de esta frase traducida china es falsa, es que las acciones de la paciente causa la propagación de virus. No es la relación de que los virus estén propagando, mientras los pacientes estén haciendo algo.

Además, si se traduce de manera literal sin agregar el contenido omitido, es fácil obstaculizar la comprensión del usuario. En la traducción de “un régimen más corto”, si no se añade “del tratamiento” ignorado en la traducción automática, no se puede entender la frase traducida en chino. En español se puede omitir el contenido respectivo con frecuencia para hacer la frase más exacta y precisa, pero en chino, suelen traducirse todas las partes de una frase. Al mismo tiempo, los adjetivos españoles suelen contener más de un significado, por lo que en la traducción deben elegir un más adecuado según el sustantivo. Si no se adapta el significado, es muy difícil para un chino entender lo que quiere transmitir, como por ejemplo “小注射” “活杆菌” “压倒性的” (los errores ya se han analizado en el apartado arriba).

#### **4.1.2 Lenguaje Coloquial**

De acuerdo con Antonio Briz Gómez (1996), el lenguaje coloquial es “el habla entre amigos, en la calle”, “Es la conversación diaria”, “Es lo oral”. Obviamente, en el ámbito médico, una ciencia médica muy rigurosa y seria, que está relacionado estrechamente con la vida y la sanidad de la población, hace falta aplicar el lenguaje escrito, profesional para garantizar la exactitud de la traducción.



Pero, después de analizar los párrafos traducidos por el iFLYTEK, se propone que en la traducción automática falta el uso de términos específicos de medicina y la expresión es muy coloquial. En el texto original se utilizan muchas veces “personas” y “usted”. El iFLYTEK traduce “personas” de manera literal como “人”, lo que en el ámbito médico hace a los pacientes o usuarios sentirse muy informal. Es mejor traducirlas como “患者 (pacientes o enfermos)”, más profesional y de estilo escrito. Además, las traducciones de “不舒服(sentirse enfermo)”, “吃药(tomar medicina)” son muy habladas. También es necesario complementar el contenido implicado del texto original, como, por ejemplo, en “las medicinas de la enfermedad” “结核病药物”, hace falta referirse cuál es la enfermedad específicamente. Desde aquí, se encuentra que la traducción de iFLYTEK no es capaz de ajustar el estilo del texto traducido según el ámbito o campo del texto original.

#### **4.1.3 Falta de correspondencia**

El contenido correspondiente entre el idioma original y el idioma es un principio básico y el objetivo principal de la traducción. Para un traductor, sea traductor automático o sea traductor humano, transmitir el significado y la intención del texto original consiste en lo más importante en la traducción. En la traducción automática realizada por iFLYTEK existen un caso de que el significado traducido no se corresponde al texto original. Solo hay una zona, donde se traduce “puede suceder varios años después” como “也可能会发生几次。你们等一下。”。 La frase traducida no tiene nada que ver con la frase española.

Aunque solo existe un error de falta de correspondencia y parece que se no puede considerar una característica de iFLYTEK, es un error grave a que deben prestar atención. Se pone este error en la lista de característica a fin de destacar la importancia de la transmisión del significado.

#### **4.1.4 Problema de forma**

La traducción automática es capaz de transformar la mayoría de los signos de español a los de chino. Es un punto muy favorable para ahorrar la edición y revisión en la etapa posterior después de la traducción. En particular en los textos médicos emitidos a los pacientes por los personales médicos, las formas aplicadas son diversas para hacer la información más accesible y aceptable. En cuanto al corpus establecido de esta tesis, las informaciones como los síntomas, el tratamiento, la prevención, las pruebas se recogen en los manuales o noticias de promoción donde hay muchas imágenes y variadas formas. Entonces, mantener la correspondencia de forma consisten en un aspecto importante para la traducción en el ámbito sanitario.

Sin embargo, el iFLYTEK no tienen capacidad de transformar signos especiales aplicados en el texto original. Para enumerar diferentes casos paralelos, en el corpus español se utiliza muchas veces “.”, pero la máquina no puede reconocer esta forma y los reemplaza

por “?”, lo que hace la forma traducida un poquito rara. Además de no poder ajustar signos especiales, en la traducción automática también aparece el mal uso de signos. Por ejemplo, se traduce “una aguja pequeña” como “小针?”. No se sabe de donde proviene esta integración y la aparición de este signo no tiene ni pies ni cabezas. Entonces, el iFLYTEK necesita mejorar la función en este aspecto. Con solucionarlo, puede tener más poderosa competencia en la traducción en el ámbito sanitario y mejorar la calidad no limitada al significado.

#### **4.1.5 Falta de contexto**

Las palabras, frases, párrafos, textos de un corpus para traducir es un conjunto, donde la persona y el estilo suelen mantener uniforme, por lo que la traducción debe permanecer uniforme también. Como se explica antes, en español se suele omitir la parte repetitiva. Si se aplica la persona usted en el texto, no hace falta escribir usted cada vez antes de una frase, es que las conjugaciones de verbos pueden reflejan la persona y la repetición no es necesaria.

Sin embargo, en chino los verbos no tienen conjugaciones, por lo que, si no se escribe la persona, no se puede saber el sujeto haciendo la frase incompleta y llevando ambigüedad. En el texto traducido por el iFLYTEK aparecen muchas zonas donde se omite la persona usted o se equivoca del sujeto. La razón radica en que las conjugaciones del verbo de usted y la tercera persona son iguales. Entonces, para solucionar este problema, la traducción necesita cambiar la traducción palabra por palabra, y conectar el contexto. Por medio de analizar el texto anterior, no solo puede ajustar la persona sino también mejorar el uso de palabras traducidas evitando el contenido redundante e innecesario. De esta manera, la traducción automática puede ser más exacta y concisa.

#### **4.2 Comentarios de métricas de evaluación de BLEU**

En el tercer capítulo se aplican las métricas de BLEU para evaluar la traducción automática realizada por BLEU. Después de analizar cada parte de la traducción, se encuentran unas ventajas y desventajas de la herramienta BLEU.

Por un lado, la herramienta BLEU tiene muchos ventajas. Se puede analizar el porcentaje de similitud entre la traducción humana de referencia y la traducción automática de manera muy rápida. Se explica el porcentaje de similitud de todo el texto y cada parte por medio de los datos, siendo muy claro y organizado. Además, se explican gráficos de barras para interpretar la similitud de cada parte, lo que facilita el análisis y la comparación posterior. Entre los marcos, se marcan las disimilitudes en color azul y similitudes en color rojo, haciendo la diferencia más obvia y evidente. En resumen, se trata de una herramienta muy útil para comparar la traducción de referencia y la traducción humana. De otra manera, podemos aprovechar esta función no limitada a comparar las traducciones, sino aplicarla en el análisis de dos textos de un mismo tema o textos estadísticos. Se puede llevar a cabo a través de estas métricas.

Por otro lado, en cuanto a la manifestación en la evaluación de la calidad de la traducción automática, desde mi punto de vista, las métricas no presentan muy bien. Se resumen sus problemas y desventajas en los siguientes aspectos, y espero que pueda mejorar la calidad y funciones la BLEU.

#### **4.2.1 No se puede distinguir los errores.**

El teorema de métrica de BLEU consiste en la comparación de la traducción de referencia y la traducción automática para encontrar las similitudes y disimilitudes. Pero las zonas marcadas en rojo no significan que sean falsas sino solo significan que hay diferencia entre la traducción humana y traducción automática. Como se señala en la figura 33 y la figura 34, entre las 146 disimilitudes no respetivas, el caso de que la traducción automática es adecuada ocupa más de 60%. Para expresar un significado hay muchas maneras de expresión. Se puede usar adjetivo, verbo, sustantivo, sudoración o adverbio de acuerdo con diferentes hábitos y diferentes preferencias, lo que no tiene la diferencia entre lo falso o correcto. El objetivo de traducción radica en la transmisión del significado del texto original y para alcanzar este objetivo existen muchísimas diversas expresiones. Diferentes traductores humanos o automáticos tienen diferentes pensamientos, así casi es imposible traducir un mismo texto a una traducción misma.

Como, por ejemplo, se traduce “感染疾病”, y la traducción no se limita a una o dos: “la infección de una enfermedad”, “infectarse de la enfermedad”, “contagiar la enfermedad”, “el contagio de una enfermedad” “infección de enfermedades”, etc. Solo con respecto a una locución, existen tantas formas de traducción, ni siquiera para una frase, para un párrafo, para un texto, para un corpus. No se puede imaginar hay cuántas maneras de combinación y ajustes de palabras. Entonces, solo por medio de comparar la similitud de texto de referencia y la traducción, no es posible distinguir los errores.

#### **4.2.2 Los parámetros no son fijos.**

Para el uso de las herramientas de AntConc o BLEU u otras herramientas semejantes en el lenguaje chino, hace falta realizar la segmentación de palabras de antemano, es que no hay espacios basados en la semántica y el significado de lenguaje. De esta manera, hay diferentes maneras para segmentar las palabras, lo que influye los parámetros en el análisis realizado por BLUE.

En una parte se han marcados disimilitudes en color ojo, pero también se marcan otra vez en otra parte, lo que se incluye en el resultado del porcentaje de la diferencia. Desde el punto de vista de teorema de funcionamiento de este tema, se separa un párrafo como palabra por palabra, así se compara de esta manera, ignorando la integridad de la frase y del texto. Existe el caso de que las palabras de una frase son mismas, pero por medio de diferentes combinaciones, expresan significados completamente diferentes. Es recomendable añadir un aspecto de considerar el orden de la frase en el funcionamiento

de sistema, lo que en mi opinión se puede elevar en algún sentido el nivel de exactitud.

#### **4.2.3 Se limita al texto de referencia.**

Para evaluar la traducción automática, la métrica BLEU toma la traducción humana de referencia. Pero desde la relación lógica no tiene razón. Por un lado, aunque la traducción de referencia se realiza por traductores profesionales o traductores bilingües, es difícil garantizar que esta traducción es cien por cien perfecta. Si existiera un error, esta herramienta no puede distinguir sino considerar que la traducción automática es falsa. Por otro lado, en cuanto a un mismo texto, no hay solo una manera de traducción o expresión. El sistema BLEU se limita al texto de referencia y toma este texto como una respuesta perfecta y única, lo que no es razonable aplicarlo en la evaluación de traducción. De todas maneras, la traducción es un proceso de pensamiento de transmitir el significado del idioma original al idioma de destino de un traductor. Debido a la diferencia de contextos culturales, niveles de estudio, y experiencias de vida, los traductores tienen diferentes maneras a fin de describir un mismo asunto.

En resumen, en base al análisis de calidad de la traducción automática marcada en color rojo por BLEU, estas métricas tienen muchas limitaciones para realizar la evaluación. No tiene un sistema lógico para distinguir los errores o disimilitudes. Además, es imposible reconocer las zonas posibles excelentes en la traducción automática, es que se toma el texto de referencia como punto de salida. Pero, aunque no hace muy bien en la evaluación de traducción, es posible aplicar BLEU para analizar y comparar los dos textos para satisfacer la necesidad de llevar al cabo el estudio o investigación de un tema o un ámbito. De todas maneras, prefiero mucho la manera de expresar el resultado de BLEU, donde los gráficos utilizados y dos datos son muy convenientes y visuales. En resumen, la BLEU es una herramienta que tienen mucha potencial para desarrollar.

### **Conclusión**

De acuerdo con los objetivos planteados en la introducción, después de la redacción de este estudio, se contesta de manera resumida:

Las características del lenguaje médico español consisten en que existen el uso de la voz pasiva, el abuso del gerundio, el solecismo, muchas terminologías, extranjerismo innecesario y falsos amigos. En el lenguaje chino médico se aplican con frecuencia la voz pasiva, expresiones fijas, y expresiones de normalización y de prudencia. Entonces, los traductores necesitan tener en cuenta los rasgos en la transmisión de los dos idiomas. Es un punto de dificultad de terminologías médicas y también es una importante competencia para ser un traductor capacitado.

Las principales métricas de evaluación automática son: BLEU (Bilingual Evaluation Understudy), METEOR (Metric for Evaluation of Translation with Explicit Ordering) y

TER (Translation Edit Rate). Los principales modelos de evaluación humana son: Modelo “Xin”, “Da”, y “Ya”, Modelo SAEJ2450 y Modelo TAUS. La evaluación automática se puede aplicar en un corpus de gran cantidad con alta eficiencia, mientras el resultado se presenta en forma de datos y gráficos, de manera muy clara. Pero las métricas BLUE no son adecuadas para distinguir errores de traducción sino para comparar la similitud de traducción y texto de referencia. La evaluación humana puede reconocer y corregir los errores de traducción, pero se limita a la cantidad de palabras. Además, en algún sentido, la evaluación humana es subjetiva y diferentes traductores pueden tener diferentes resultados de la calidad de una misma traducción.

La calidad de la traducción de texto de tuberculosis realizada por iFLYTEK puede satisfacer la necesidad de pacientes de manera general. Pero también existen problemas que puede mejorar la empresa. Por un lado, la técnica aplicada es la traducción literal, lo que hace la traducción redundante y menos concisa. Además, se usa el lenguaje coloquial, bajando la exactitud del texto médico. Por otro lado, dispone de problemas de formas y conexión de contexto.

Esta tesis está dedicada al análisis del traductor iFLYTEK en la traducción en el ámbito sanitario. Pero debido a las palabras y energía, se concentra a la enfermedad de tuberculosis, lo que hace este estudio limitado. Hace falta más investigación y estudio para explotar la aplicación de este traductor en más especialidades del campo médico.

## Bibliografía

- Abril Martí, M.I. (2006). *La interpretación en los servicios públicos: Caracterización como género, contextualización y modelos de formación. Hacia unas bases para diseño curricular*. Universidad de Granada
- Albir, A. H. (2001). *Traducción y traductología / Translation and translation studies: Introducción a la traductología / Introduction to Translation*. Catedra Ediciones.
- Aledixandre Benavent, R. y Amador Iscla, A. (2001) “*Problemas del lenguaje médico actual (I): Extranjerismos y falsos amigos*” en Panace@, 10(3) Valencia: Tremédica.144-149
- Aledixandre Benavent, R. y Amador Iscla, A. (2001) “*Problemas del lenguaje médico actual (II): Abreviaturas y epónimos*” en Panace@, 10(4) Valencia: Tremédica.170-176
- Aledixandre Benavent, R. y Amador Iscla, A. (2002) “*Problemas del lenguaje médico actual (III): Gramáticas y estilo*” en Panace@, 11(1) Valencia: Tremédica.18-23
- Benavant. R. A. , Bueno Cañigral F. J., Cogollos L. C., Rafael Aleixandre-Benavent, (2017), *Características del lenguaje médico actual en los artículos científicos*, Educ Med.

- Bernal Chávez, J. A., & Hincapié Moreno, D. A. (2018). *Lingüística de corpus*.
- Bianca Vitalaru, (2022), *PPT, Tipología textual-sanitario*, Traducción especializada en el ámbito sanitario, Universidad de Alcalá
- Bianca Vitalaru, (2022), *PPT, traducción y documentación*. rev.pdf. p.10, Traducción especializada en el ámbito sanitario, Universidad de Alcalá
- Briva-Iglesias, V. (2021). *Traducción humana vs. traducción automática: análisis contrastivo e implicaciones para la aplicación de la traducción automática en traducción jurídica*. Mutatis Mutandis. Revista Latinoamericana de Traducción, 14(2), 571-600.
- Buthmann, S. (2013). Un concepto ampliado de traducción especializada. *AnMal Electrónica*, 34, 265-269.
- Cabré, M. T. (1993). *La terminología. Teoría, metodología y aplicaciones*. Barcelona: Editorial Antártida/Empúries.
- Colado, S. C., Real, E., Jiménez, D., Pujante, D., & Cortijo, A. (2001). *El origen de los errores en traducción*. *Écrire, traduire et représenter la fête*, 813-822.
- Diéguez M., M. I. ., & Cabrera P., I. . (1997). *Traducción humana o traducción automática: variables para una toma de decisiones eficiente*. *Onomázein*, (2), 227–242.
- Esteban Hernández, S. (2020). *Evaluación de la calidad de la traducción de motores de traducción automática neuronal en textos del ámbito jurídico*.
- Fernández-Parra, M. (2009). *The workflow of computer-assisted translation tools in specialised translation*. In *Reconceptualizing LSP: Online proceedings of the XVII European LSP Symposium*.
- Garcés, C. V., & Taibi, M. (2006). *El grupo FITISPos, formación e investigación en traducción e interpretación en los servicios públicos*. *Revista española de lingüística aplicada*, (1), 205-216.
- Gómez, A. B. (1996). *El español coloquial: situación y uso* (Vol. 38). Arco Libros.
- Hanlon, J. J., (2014). *La filosofía de la salud pública*. *Revista Cubana de Salud Pública*, 40(1), . ISSN: 0864-3466.
- Hortal Alonso, Augusto (2002): *Ética general de las profesiones*. Bilbao: Desclée de Brouwer.

- Hu, K., & Cadwell, P. (2016). *A comparative study of post-editing guidelines*. In Proceedings of the 19th annual conference of the European association for machine translation (pp. 34206-353).
- Laguna, J. Y., & Cuñat, V. A. (2011). *Diccionario de siglas médicas*. SEDOM.
- Lázaro Gutiérrez, R. (2009). *Dilemas éticos del traductor/intérprete en los servicios públicos*. Translation, Interpreting, Social Activism, sin paginar.
- Marié, E. (1998). *Compendio de medicina china: fundamentos, teoría y práctica* (Vol. 180). Edaf.
- Martínez de Cañas Pagola, L. C. (2017). *Análisis de la Formación en Traducción médico-sanitaria en España: una primera aproximación*. Universitat Jaume I
- Mayor, A., Alegria, I., de Ilarraza, A. D., Labaka, G., Lersundi, M., & Sarasola, K. (2009). *Evaluación de un sistema de traducción automática basado en reglas o por qué BLEU sólo sirve para lo que sirve*. Procesamiento del Lenguaje Natural, (43), 197-205.
- Mercedes, P. H. (2002). *En torno a la traducción automática*. Cervantes, 1(2), 101-117.
- Montalt, V., y M. González Davies (2007): *Medical translation step by step: learning by drafting*. Manchester: St. Jerome
- Muñoz Miquel, Ana, (2016), *La traducción médica como especialidad académica: Algunos rasgos definitorios*. Hermeneus. Revista de la Facultad de Traducción e Interpretación de Soria Núm. 18 Pág. 235-267
- Muñoz-Miquel, Ana. (2014). *El perfil del traductor médico: análisis y descripción de competencias específicas para su formación*. Tesis doctoral. Castelló de la Plana: Universitat Jaume I.
- Muñoz-Miquel, Ana. "El perfil del traductor médico: diseño de un estudio de corte socioprofesional." Panacea. X.30 (2009): 157-167
- Nagao, M. (1984). *A framework of a mechanical translation between Japanese and English by analogy principle*. Artificial and human intelligence, 351-354.
- Natalia Maravall, (2022), Traducción especializada en el ámbito sanitario, UD-Maravall-TS, PPT, Traducción especializada en el ámbito sanitario, Universidad de Alcalá.
- Navarro, F. A. (1996). *Traducción y lenguaje en medicina*. Fundación Dr. Antonio Esteve.
- Obushikhina Danilova, Veronika. et al. (2021). *La traducción automática y posesión en el*

*ámbito de los servicios públicos (RU-ES)*. Trabajo Fin de Máster. Universidad de Alcalá.

- Pöschhacker, F. (2004). *Introducing Interpreting Studies*. Londres/Nueva York: Routledge.
- Reyes. J. M., Vázquez. R. R., Arias, S. M., (2015), *Educación Diabetológica y Ciudadanos en enfermería en las personas con diabetes en el ámbito extrahospitalario*, Grupo Enfermería de SEMES Diabetes Andalucía Sociedad Española de Medicina de Urgencias y Emergencias(SEMES) Andalucía.
- Sánchez Ramos, & Rico Pérez. (2020). *Traducción automática : conceptos clave, procesos de evaluación y técnicas de posesición*. Comares.
- Trujillos-Yébenes, L., & Muñoz-Miquel, A. (2022). *La traducción automática y la posesición en el ámbito médico*. Tradumática: tecnologías de la traducción, (20).
- Valero-Garcés, C. et al. (Ed.). (2014). *(Re)considerando ética e ideología en situaciones de conflicto/ (Re)visiting ethics and ideology in situations of conflict*. Alcalá de Henares: Servicios de publicaciones de la Universidad de Alcalá.
- Ye Lu,(2020), *La terminología aplicada a la traducción médica: elaboración de un glosario bilingüe chino-español*, TFM, Máster Universitario en Comunicación Intercultural, Interpretación y Traducción en los Servicios Públicos
- I. ZAPATA, “*Un recorrido por algunos principios de la Traducción Automática*” en Enseñanza de lenguas y traducción con ordenadores, ed. G. Ruipérez, Madrid:Ediciones pedagógicas, 1995, 87-126.
- II.  
蔡毅, (2003), *关于翻译理论中的几个基本概念问题*, 中国俄语教学, 第22卷第1期 (Cai. Y., (2003), *Sobre varias cuestiones conceptuales básicas en la teoría de la traducción*, Enseñanza del ruso en China, Volumen 22, Número 1)
- 曾江霞. (2020). *多模态大数据语境下科技术语翻译标准分析——以新冠肺炎和新冠病毒术语翻译为例*. 中国科技术语, 22(5), 33. (Zeng Jiangxia. (2020). *Análisis de estándares de traducción de términos científicos y técnicos en el contexto de big data multimodal: ejemplos de traducción de los términos neotripanosomia y coronavirus*. Términos científicos y técnicos chinos, 22(5), 33.)
- 陈相 (2008), *面向生物医学领域的双语对齐技术研究*, 大连理工大学 (Chen, Xiang (2008), *Investigación sobre tecnología de alineación bilingüe para el ámbito biomédico*, Universidad Tecnológica de Dalian.)
- 范守义. 1987. *模糊数学与译文评价*. 中国翻译 (Fan Shouyi. 1987. *Matemáticas difusas y evaluación de la traducción*. Traducción china)



- 冯志伟. (2010). *基于语料库的机器翻译系统*. 术语标准化与信息技术, 计算机语言学. 第 1 期. (Feng Z. w. (2010). *Sistema de traducción automática basado en corpus*. Estandarización terminológica y tecnología de la información, Lingüística informática. No. 1)
- 郭子凯, 陈晨, & 刘曼玲. (2022). 译者主体性视角下医学学术文本“机器翻译+译后编辑”中的策略研究. 中国医药导报, 19(5), 192-196. (Guo Zikai, Chen Chen, & Liu Manling. (2022). *Investigación de estrategias en "traducción automática + postedición" de textos académicos médicos desde la perspectiva de la subjetividad del traductor*. China Medical Journal, 19(5), 192-196.)
- 李亚超,熊德意 & 张民.(2018). *神经机器翻译综述*. 计算机学报(12),2734-2755. (Li Yachao,Xiong Deyi y Zhang Min. (2018). *Una revisión de la traducción automática neuronal*. Revista de ciencias de la computación (12), 2734-2755.)
- 李业刚,黄河燕,史树敏,冯冲 & 苏超.(2015). *多策略机器翻译研究综述*. 中文信息学报 (02),1-9+23. (Li Yegang, Huang Heyan, Shi Shumin, Feng Chong y Su Chao. (2015). *Una revisión de la investigación en traducción automática multi-estrategia*. Revista china de informática (02), 1-9+23.)
- 刘洪涛, (2020), *基于生物医学语料的神经机器翻译系统研究*, 吉林大学 (Liu, Hongtao, (2020), *Investigación sobre un sistema neural de traducción automática basado en un corpus biomédico*, Universidad de Jilin.)
- 秦颖. (2015). *翻译质量自动评价研究综述*. 计算机应用研究, 32(2), 326-329. (Qin Ying. (2015). *Una revisión de la investigación sobre la evaluación automática de la calidad de la traducción*. *Investigación sobre aplicaciones informáticas*, 32(2), 326-329.)
- 盛诗梦, 丁皓, 徐欣茹, 丁思吉, & 夏冬阳. (2022). *注意力机制在医学上的应用综述*. (Sheng Shimeng, Ding Hao, Xu Xinru, Ding Siji, & Xia Dongyang. (2022). *Una revisión de los mecanismos atencionales para aplicaciones médicas*. )
- 孙逸群, & 周敏康. (2017). *机器翻译质量综合评价方法研究*. 中国科技翻译, 30(2), 20-24.(Sun, Yijun, y Zhou, M.K.. (2017). *Investigación sobre el método de evaluación integral de la calidad de la traducción automática*. Chinese Scientific and Technical Translation, 30(2), 20-24.)
- 王茜, (2007), *关于医院管理中医学语言质量管理的思考*, 中国医学伦理学, 第 20 卷第 4 期 (Qian, W., (2007), *Reflexiones sobre la gestión de calidad del lenguaje de la medicina tradicional china en la gestión hospitalaria*, Ética médica china, volumen 20, número 4)
- 武光军. 2006. *《翻译质量评估: 论辩理论模式》评介*[J].外语研究 (Wu, Guangjun. 2006. *Revisión de la evaluación de la calidad de la traducción: un modelo teórico de*

*argumentación [J]. Estudios de Lenguas Extranjeras)*

郁青青, 石冯珺, (2012), *机器翻译系统的历史与研究现状分析*, A, 淮海工学院学报(人文社会科学版), 第 10 卷第 12 期 (Qingqing Yu, Jun Shi, (2012), *Análisis de la historia y el estado de la investigación de los sistemas de traducción automática*, A, Journal of Huaihai Institute of Technology (Humanities and Social Sciences Edition), Vol. 10, No. 12

袁小于.(2011). *基于规则的机器翻译技术综述*. 重庆文理学院学报(自然科学版)(03),56-59.doi:10.15998/j.cnki.issn1673-8012.2011.03.016. (Yuan Xiaoxiao. (2011). *Una revisión de las técnicas de traducción automática basadas en reglas*. Journal of Chongqing College of Arts and Sciences (Natural Science Edition) (03), 56-59. doi:10.15998/j.cnki.issn1673-8012.2011.03.016.)

张邵源, (2021), *面向生物医学领域的神经机器翻译领域适用方法研究*, 华南理工大学 (Zhang Shaoyuan, (2021), *Investigación sobre métodos aplicados en el campo de traducción automática neonatal en el ámbito de la biomedicina*, Universidad Tecnológica del Sur de China.)

张涛, (2011), *面向特定领域的统计机器翻译研究与应用*, 昆明理工大学 (Zhang, Tao, (2011), *Investigación y aplicación de la traducción automática estadística para un dominio específico*, Universidad Tecnológica de Kunming.)

张霄军. (2007). *翻译质量量化评价研究综述*. 外语研究, (4), 80-84. (Zhang Xiaojun. (2007). *Revisión de la investigación sobre la evaluación cuantitativa de la calidad de la traducción*. Estudios de Lenguas Extranjeras, (4), 80-84.)

朱德熙,《语法讲义》,北京:商务印书馆,1982. p25 (Zhu. D. X, *Conferencias de gramática*, Commercial Imprenta de Beijing, 1982. p25)

### **Herramientas utilizadas**

ATAN: <https://fanyi.atman360.com/>

ATMAN: <https://fanyi.atman360.com/index#indexJoinTrial>

BLUE: <https://www.letsmt.eu/Bleu.aspx>

Diccionario: Real Academia Española

Diccionario: Real Academia Nacional de Medicina de España

Enciclopedia Baidu

Gobierno de España, Ministerio de Industria, Comercios, y Turismo:

<https://www.tourspain.es/eu-es/ConocimientoTuristico/FichasEjecutivas/Ficha%20ejecutiva%20China%20Mayo%202020.pdf>

iFLYTEK: <https://www.xfyun.cn/?ch=xfow>.

Instituto Nacional de Estadística: [https://www.ine.es/prensa/pad\\_2022\\_p.pdf](https://www.ine.es/prensa/pad_2022_p.pdf)

La Prueba de Acreditación de China para Traductores e Intérpretes, 全国翻译专业资格考试综合服务平台(CATTI): <http://www.catticenter.com/cattixbyy>

Medine Puls: <https://medlineplus.gov/spanish/ency/article/001933.htm>

Nimdzi Language Technology Atlas 2021:  
<https://www.nimdzi.com/nimdzi-language-technology-atlas-2021/>

Notepad ++ : <https://notepad-plus-plus.org/downloads/v8.5.2/>

Red de Revistas Médicas Chinas (China Medical Journals Network):  
<https://www.medjournals.cn/index.do>

ROST: <https://pan.baidu.com/s/1TPMhpEOAT--BfNPmRdTzqA>

SKOS: Nomenclatura de Ciencia y Tecnología de la UNESCO  
(<https://skos.um.es/unesco6/32/html>)

SYSTRAN:

<https://www.systransoft.com/es/systran/tecnologia/que-es-la-traduccion-automatica/>

复旦大学实验动物科学部( Departamento de Ciencia de Animales de Laboratorio, Universidad de Fudan): <http://labanimal.fudan.edu.cn/26586/list.htm>, última citada: 2023

## **Anexos**

### **Anexo 1: Texto original en chino (de referencia)**

“什么是潜伏结核感染(LTBI)?

潜伏结核感染或 LTBI 系指您体内呈“休眠”状态的存活结核病菌。潜伏结核感染患者并无任何症状、不会感到身体不适，也不会将结核病菌传播给其他人。然而，若患者的身体免疫功能低，这些结核病菌将会“苏醒”并开始繁殖。随后，活动期结核病菌会使患者

生病。活动期结核病系指人体内的结核病菌仍然存活并已“苏醒”。活动期结核病患者会有严重咳嗽、夜间盗汗、疲劳和体重减轻等症状。而且，这些患者会把结核病菌传播给其他人。结核病菌主要影响肺脏，但也可对人体其它器官产生影响。若不加以治疗，结核病可导致死亡。”(Condado de los Angeles Salud Pública:ww.publichealth.lacounty.gov/tb)

“结核病菌如何传染？

当活动期结核病患者咳嗽、打喷嚏、发笑、讲话或唱歌时，结核病菌便会在空气中传播。若您吸入带有结核病菌的空气时，便会受感染。然而，您不会经由衣物、盘子、食物、床上用品或与活动期结核病人的实际接触而患结核病。”(Condado de los Angeles Salud Pública:ww.publichealth.lacounty.gov/tb)

“对潜伏结核感染有哪些治疗方法？

若您受结核病菌感染，则会患结核病。您可能会立即发病，也可能于几年之后发病。对潜伏结核感染可采用药物治疗；最常用的药物为雷米封(INH)。服药期必须持续六至九个月。(儿童及受艾滋病毒感染(HIV)者需要更长的服药期；至少应为九个月。)”(Condado de los Angeles Salud Pública:ww.publichealth.lacounty.gov/tb)

“如果我并未感到身体不适，为何仍需服药？

如果您不服用潜伏结核感染治疗药物，这种感染则会发展为活动期结核病。开始服药后，许多结核病菌将被杀灭，但仍会有一些结核病菌在您体内存活很长时间。只要体内有结核病菌存活，这些病菌便会“苏醒”、繁殖并使您生病。若您服用结核病治疗药物，则必须按照医生或医疗保健服务提供者所开的处方服用全部药物；这十分重要。”(Condado de los Angeles Salud Pública:ww.publichealth.lacounty.gov/tb)

“结核病感染检测有两种类型。

您的医疗服务提供者可能会推荐结核血检，因为它可以在一次就诊中完成，而且如果您过去曾接种过结核病疫苗，那么它更准确。血液检测测量您的免疫系统对引起结核病的病菌的反应。在进行皮试时，要用一根小针将一些检测材料放在皮肤下。您需要在 2-3 天内回到您的医疗保健医生那里，看看是否有反应。”(CDC: www.cd.gov/thinktesttreattb)

“皮试结果若呈阳性会怎样？

阳性皮试结果通常表示您已经受到受结核病菌感染，但未必表示您已经患有活动期结核病。您需接受其它测试，例如胸部 X 光透视或痰（痰液）采样等，藉以查明您是否患有活动期结核病。”(Condado de los Angeles Salud Pública:ww.publichealth.lacounty.gov/tb)

“应该于何时接受皮试？

若存在以下情况，您应接受测试：

- 存在活动期结核病症状
- 曾经与活动期结核病患者有长时间接触
- 患有会削弱人体免疫系统的病症，例如艾滋病毒感染(HIV)或糖尿病等。
- 曾经注射过非法药品
- 最近从某个结核病发病率高的地区来到美国，这些地区包括拉丁美洲、加勒比海地区、非洲、亚洲或东欧地区。
- 曾在护养院、无家可归者救济所、监狱或其它集体环境中工作或住宿。”(Condado de los

Angeles Salud Pública:[www.publichealth.lacounty.gov/tb](http://www.publichealth.lacounty.gov/tb))

“潜伏结核感染(LTBI)治疗药物可能会产生哪些副作用？

此类药物的副作用十分罕见，但可能包括：

- 恶心
- 呕吐
- 尿液呈深色
- 皮疹
- 皮肤和（或）眼睛呈黄色
- 食欲不振
- 双手和（或）双足麻木” (Condado de los Angeles Salud Pública:[www.publichealth.lacounty.gov/tb](http://www.publichealth.lacounty.gov/tb))

“我没觉得身体不适，如何判断被感染呢？

结核感染不会有任何症状，所以我们会通过特殊检查来诊断。您可以在当地结核服务机构做检查，也可以得到他们的相关帮助。检查包括：

- 通过  $\gamma$  干扰素释放试验血检 (IGRA 或 QuantiFERON TB Gold Plus®) 可判断您是否感染了结核病。
- 通过结核菌素皮试（又称作 TST 或结核菌素试验），在前臂进行量注射。注射 2 至 3 天之后再次检查，根据身体反应（红肿）判断是否被感染结核病。
- 通过胸部 X 光可判断肺部是否有结核。有时 X 光还可查出之前的结核，即使您并不知道之前患过结核病。” (NSW Government:[www.health.nsw.gov.au/Infectious/tuberculosis/](http://www.health.nsw.gov.au/Infectious/tuberculosis/))

“卡介苗疫苗(BCG)是什么？

卡介苗疫苗是一种活性疫苗，可有限预防结核病。若接触到结核病病菌，卡介苗疫苗不能保护您免受感染，但却极有可能把感染变成轻微的局部感染，而不是威胁生命的重病。卡介苗疫苗需要 6 至 12 周时间来产生免疫力保护，预防结核病的免疫能力可能只有 50-60%，有些人体内的疫苗会在 5-15 年间失去效力。” (NSW Government:<https://www.health.nsw.gov.au/Infectious/tuberculosis/Documents/Language/bcg-information-spa.pdf>)

“卡介苗疫苗和其他活性疫苗。

前往海外旅游而需要接种超过一种活性疫苗的人士，应该：a) 在同一天接种所有活性疫苗，或 b) 接种卡介苗疫苗和其他活性疫苗之间等候最少四星期。这样做是让身体的免疫 (防卫) 系统最佳的机会，提供所需的免疫力(保护)。其他的活性疫苗包括：口服伤寒疫苗，黄热病疫苗，腮腺炎、麻疹、德国麻疹混合疫苗。卡介苗疫苗和口服小儿麻痹疫苗可在四星期内接种。”(NSW Government:<https://www.health.nsw.gov.au/Infectious/tuberculosis/Documents/Language/bcg-information-spa.pdf>)

## **Anexo 2: Textos originales en español**

“¿Qué es la infección de la TB (LTBI)?

La infección de la TB o LTBI, significa que los gérmenes de la TB están vivos, pero

“dormidos” en su cuerpo. Si tiene la infección de la TB, usted no tendrá síntomas, es posible que no se sienta enfermo(a) y no puede contagiar los gérmenes de la TB a otras personas. Sin embargo, si su sistema inmunológico se debilita, estos gérmenes de la TB pueden “despertar” y multiplicarse. Y entonces, puede enfermarse con TB activa. Si usted tiene la enfermedad de la TB, presentará síntomas como tos fuerte, sudor por la noche, cansancio y pérdida de peso. Además, puede contagiar a otras personas con los gérmenes de la TB. La TB afecta especialmente a los pulmones, pero también puede afectar otras partes del cuerpo. Si no se trata, la enfermedad de la TB puede causar la muerte.”(Condado de los Angeles Salud Pública: [ww.publichealth.lacounty.gov/tb](http://ww.publichealth.lacounty.gov/tb))

“¿Cómo se contagian las personas?

Los gérmenes de la TB se propagan en el aire cuando una persona que está enferma con TB activa tose, estornuda, se ríe, habla o canta. Si respira aire que tiene gérmenes de la TB, puede infectarse. Sin embargo, usted no puede contraer la TB a través de la ropa, los platos, la comida, la ropa de cama o el contacto físico con alguien que tenga TB activa.” (Condado de los Angeles Salud Pública: [ww.publichealth.lacounty.gov/tb](http://ww.publichealth.lacounty.gov/tb))

“¿Cuál es el tratamiento para la infección de la TB (LTBI)?

Si se infectó con los gérmenes de la TB, puede enfermarse de la TB. Esto puede suceder de inmediato o varios años después. La LTBI puede tratarse con medicinas. La más usual es la isoniazida (INH). Debe tomarse esta medicina durante seis a nueve meses. También hay un régimen más corto, que combina dos medicinas que se toman semanalmente durante 12 semanas.”(Condado de los Angeles Salud Pública: [ww.publichealth.lacounty.gov/tb](http://ww.publichealth.lacounty.gov/tb))

“¿Por qué debo tomar medicina si no me siento enfermo(a)?

Puede desarrollar la enfermedad de la TB si no toma la medicina para su infección de la TB. Muchos gérmenes de la TB mueren después de empezar a tomar la medicina, pero algunos permanecen vivos en el cuerpo por un largo tiempo. Mientras usted todavía tenga gérmenes de TB en su cuerpo, pueden “despertar”, multiplicarse y hacer que se enferme. Si está tomando medicinas para la TB, es importante que termine su medicina, de la forma en que lo haya indicado el médico o profesional de la salud.” (Condado de los Angeles Salud Pública: [ww.publichealth.lacounty.gov/tb](http://ww.publichealth.lacounty.gov/tb))

“Hay dos tipos de pruebas para detectar la infección de tuberculosis.

Su proveedor de atención médica podría recomendar el análisis de sangre de detección de tuberculosis porque se puede hacer en una visita y es más preciso si usted se ha vacunado contra la tuberculosis en el pasado. El análisis de sangre mide cómo reacciona su sistema inmunitario ante los microbios que causan la tuberculosis. En la prueba cutánea de la tuberculina, se usa una aguja pequeña para poner algo de material de análisis debajo de la piel. Usted deberá volver a su proveedor de atención médica en 2 o 3 días para ver si hay una reacción.” (CDC: [www.cdc.gov/thinktesttreattb](http://www.cdc.gov/thinktesttreattb))

“¿Qué pasa si el resultado de la prueba de la piel es positivo?

Un resultado positivo de la prueba de la piel por lo general significa que usted está infectado con gérmenes de la TB, pero no significa precisamente que usted tiene la enfermedad de la TB. Se necesitan otras pruebas, como una radiografía de tórax o una muestra de esputo (flema), para averiguar si usted tiene la enfermedad de la TB.” (Condado de los Angeles Salud Pública: [ww.publichealth.lacounty.gov/tb](http://ww.publichealth.lacounty.gov/tb))

“¿Cuándo debo hacerme la prueba de la piel?

Debe hacerse la prueba de la TB si:

- Tiene síntomas de la enfermedad de la TB
- Ha pasado mucho tiempo con alguien que tiene enfermedad de la TB
- Tiene una condición que debilite su sistema inmunológico, como el VIH o la diabetes
- Se ha inyectado drogas ilegales
- Ha venido a Estados Unidos recién de una zona con muchos casos de TB, como América Latina, Caribe, África, Asia o Europa del Este
- Ha trabajado o ha pasado tiempo en un hogar de ancianos, un refugio para desamparados, una prisión, u otro centro para grupos de personas.” (Condado de los Angeles Salud Pública: [ww.publichealth.lacounty.gov/tb](http://ww.publichealth.lacounty.gov/tb))

“¿Cuáles son algunos posibles efectos secundarios de las medicinas de la enfermedad de la TB?

Los efectos secundarios no son comunes con estas pastillas, pero pueden incluir:

- Náuseas
- Vómito
- Orina oscura
- Sarpullido en la piel
- Piel u ojos amarillentos
- Pérdida del apetito
- Adormecimiento de manos o pies” (Condado de los Angeles Salud Pública: [ww.publichealth.lacounty.gov/tb](http://ww.publichealth.lacounty.gov/tb))

“No me siento enfermo, ¿cómo puedo saber si he sido contagiado?

La infección de TB no causa ningún síntoma, por lo cual tenemos que realizar pruebas especiales para diagnosticar una infección de TB.

Las pruebas se pueden hacer o coordinar por medio del Servicio de TB de su localidad. Estas pruebas incluyen:

- Un examen de sangre llamado ensayo de liberación de interferón gamma (IGRA o QuantiFERON TB Gold Plus®) puede demostrar si usted se ha contagiado con TB.
- Una prueba cutánea de tuberculina (conocida también como TST o prueba Mantoux), es una pequeña inyección colocada justo bajo la piel de la parte interna del antebrazo. Usted deberá regresar entre 2 y 3 días después de efectuada la prueba para poder evaluar cualquier reacción (hinchazón y enrojecimiento) que indique una infección de TB.” (NSW Government: [www.health.nsw.gov.au/Infectious/tuberculosis/](http://www.health.nsw.gov.au/Infectious/tuberculosis/))

“Qué es la Vacuna BCG?

La vacuna BCG es una vacuna de bacilo vivo que brinda cierto grado de protección contra la tuberculosis. La vacuna BCG no evita que la persona resulte infectada si se ve expuesta a las bacterias (gérmenes) de la tuberculosis, pero aumenta en mucho las probabilidades de que dicha persona contraiga una infección leve y localizada, en lugar de una enfermedad grave y posiblemente fatal. La vacuna BCG puede tardar entre 6 y 12 semanas antes de producir su efecto inmune (protector). La vacunación BCG puede solamente brindar el 50-60% de inmunidad contra la tuberculosis y en algunas personas la vacuna pierde su eficacia con el tiempo, a veces en un plazo de 5-15 años.” (NSW Government: <https://www.health.nsw.gov.au/Infectious/tuberculosis/Documents/Language/bcg-information-spa.pdf>)

“La BCG y otras vacunas de bacilo/virus vivo.

Las personas que viajen al extranjero y que necesiten más de una vacuna de bacilo/virus vivo antes de viajar, debería) recibir todas las vacunas de bacilo/virus vivo el mismo día, o b) esperar por lo menos 4 semanas entre la BCG y las otras vacunas de bacilo/virus vivo. Con ello se le brinda al sistema inmune (de defensa) del cuerpo la mayor oportunidad de ofrecer la inmunidad (protección) necesaria. Otras vacunas de bacilo/virus vivo son: contra la tifoidea oral, fiebre amarilla y paperas, sarampión y rubéola (MMR). La BCG y la antipoliomielítica oral se pueden administrar dentro del período de 4 semanas.” (NSW Government: <https://www.health.nsw.gov.au/Infectious/tuberculosis/Documents/Language/bcg-information-spa.pdf>, )

### **Anexo 3: Textos traducidos al chino por iFLYTEK**

什么是结核病感染？

结核病或 LTBI 感染意味着结核病细菌是活的，但在您的体内“沉睡”。如果您感染了结核病，您就不会有任何症状，您可能不会感到不适，也不会将结核病病菌传染给其他人。然而，如果您的免疫系统减弱，这些结核病病菌可能会“苏醒”并繁殖。然后，您可能会患上活动性结核病。如果您患有结核病，您会出现剧烈咳嗽、夜间出汗、疲劳和体重减轻等症状。此外，它还会将结核病病菌传染给他人。结核病特别影响肺部，但也会影响身体的其他部位。如果不治疗，结核病可能会导致死亡。

人是怎么传染的？

结核病病菌在空气中传播时，患有活动性结核病的人咳嗽，打喷嚏，大笑，说话 或者唱歌。如果您呼吸的空气中有结核病病菌，您可能会被感染。然而，您不能通过衣服、盘子、食物、床上用品或与患有活动性结核病的人的身体接触而感染结核病。

结核病感染的治疗方法是什么？

如果您感染了结核病病菌，您可能会患上结核病。这可能会立即发生，也可能发生几次？你们等一下。LTBI 可以用药物治疗。最常见的是异烟肼（INH）。这种药应该服用六到九个月。还有一种较短的方案，结合每周服用 12 周的两种药物。

如果我不觉得不舒服，为什么要吃药？



如果您不服用治疗结核病的药物，您可能会患上结核病。许多结核病病菌在开始服药后死亡，但也有一些在体内存活了很长时间。只要您体内还有结核病病菌，它们就会“苏醒”，繁殖并使您生病。如果您正在服用结核病药物，重要的是按照医生或卫生专业人员的指示完成您的药物治疗。

有两种检测结核病感染的方法。您的医疗保健提供者可能会建议分析 结核病检测血液，因为它可以在一次访问中完成 如果您过去接种过结核病疫苗，那就更准确了。血液测试测量您的免疫系统对引起结核病的微生物的反应。在结核菌素皮试中，用小针？在皮肤下放一些分析材料。您必须在 2 或 3 天内返回您的医疗保健提供者，看看是否有反应。

如果皮肤测试结果呈阳性怎么办？

皮肤测试的阳性结果通常意味着 您感染了结核病病菌，但这并不完全意味着您患有结核病。需要其他检查，如胸部 X 光或痰（痰）样本，以确定您是否患有结核病。

我应该什么时候做皮肤测试？

如果出现以下情况，应进行结核病检测：

- ？他有结核病的症状。
- ？他花了很多时间和一个患有结核病的人在一起。
- ？有条件吗？削弱您的免疫系统，如艾滋病毒或糖尿病
- ？非法药物已被注射
- ？他刚从拉丁美洲、加勒比、非洲、亚洲或东欧等结核病高发地区来到美国。
- ？他曾在养老院、无家可归者收容所、监狱或其他人群中心工作或度过一段时间。

结核病药物可能有哪些副作用？

这些药片的副作用并不常见，但它们可能包括：

- ？恶心。
- ？呕吐。
- ？深色尿液
- ？皮肤上的皮疹
- ？皮肤或眼睛发黄
- ？食欲不振
- ？手或脚麻木

我不觉得恶心，？我怎么知道我是不是被传染了？

结核病感染不会引起任何症状，因此我们必须进行特殊测试来诊断结核病感染。

可以通过当地的结核病服务机构进行或协调测试。这些测试包括：

- ？一项名为  $\gamma$  干扰素释放试验（IGRA 或 QuantiFERON TB Gold Plus？）的血液检查可以证明您是否感染了结核病。
- ？结核菌素皮肤试验（也称为 TST 或 Mantoux 试验）是一种小注射，正好放置在前臂内侧的皮肤下。您必须在测试后 2 至 3 天内返回，以便能够评估任何表明结核病感染的反应（肿胀和发红）。

卡介苗是什么？

卡介苗是一种活杆菌疫苗，可提供一定程度的结核病保护。卡介苗并不能防止一个人在接触结核病细菌（病菌）时被感染，但它会大大增加一个人感染轻微和局部感染的可能性，而不是一种压倒性的、可能致命的疾病。卡介苗可能需要 6 至 12 周才能产生免疫（保护）作用。卡介苗疫苗只能提供 50-60%的结核病免疫力，在某些人中，疫苗会随着时间的推移而失效，有时会在 5-15 天内失效。

卡介苗和其他活杆菌/病毒疫苗。

出国旅行并在旅行前需要一种以上活杆菌/病毒疫苗的人，应（a）在同一天接种所有活杆菌/病毒疫苗，或（b）在卡介苗和其他活杆菌/病毒疫苗之间至少等待 4 周。这为身体的免疫（防御）系统提供了更大的机会来提供必要的免疫（保护）。其他细菌/病毒疫苗包括：口服伤寒、黄热病和腮腺炎、麻疹和风疹（MMR）疫苗。卡介苗和口服抗脊髓灰质炎药物可在 4 周内服用。