

# El Inglés y el Español en la Lingüística Computacional: **Traducción Automática**

Felipe Sánchez Martínez y **Miquel Esplà Gomis**

Dep. de Llenguatges i Sistemes Informàtics  
Universitat d'Alacant

Curso 2020/21

Profesor: Miquel Esplà Gomis

---

Reserva tutoría: UACloud→Tutorías

Horario: Disponible al hacer la reserva

Despacho: Escuela Politécnica Superior IV, módulo central,  
2ª planta, laboratorio de investigación 2

---

Profesor: Miquel Esplà Gomis

---

Reserva tutoría: UACloud→Tutorías

Horario: Disponible al hacer la reserva

Despacho: Escuela Politécnica Superior IV, módulo central,  
2ª planta, laboratorio de investigación 2

---

Sesiones remotas: [meet.google.com/jfj-atuj-ohz](https://meet.google.com/jfj-atuj-ohz)

Clases participativas: resolución en grupo de cuestiones y problemas

---

Profesor: Miquel Esplà Gomis

---

Reserva tutoría: UACloud→Tutorías

Horario: Disponible al hacer la reserva

Despacho: Escuela Politécnica Superior IV, módulo central,  
2ª planta, laboratorio de investigación 2

---

Sesiones remotas: [meet.google.com/jfj-atuj-ohz](https://meet.google.com/jfj-atuj-ohz)

Clases participativas: resolución en grupo de cuestiones y problemas

---

Evaluación: **25 %** de la nota final

Entrega trabajo práctico

**31/01/2021**; UACload→Evaluación

Este bloque se denomina *traducción automática*

- ¿Qué creéis que estudiaremos?

Este bloque se denomina *traducción automática*

- ¿Qué creéis que estudiaremos?
- ¿Qué disciplinas o conceptos creéis que están involucrados?

Este bloque se denomina *traducción automática*

- ¿Qué creéis que estudiaremos?
- ¿Qué disciplinas o conceptos creéis que están involucrados?
- ¿Qué idiomas habláis, escribís o traducís? ¿Con que nivel de soltura?

Este bloque se denomina *traducción automática*

- ¿Qué creéis que estudiaremos?
- ¿Qué disciplinas o conceptos creéis que están involucrados?
- ¿Qué idiomas habláis, escribís o traducís? ¿Con que nivel de soltura?
- ¿Quiénes habéis estudiado el *Grado en Traducción e Interpretación*?



## 1 Traducción automática

- Aplicaciones de la traducción automática
- Automatización del proceso de traducción

## 2 Evaluación de la traducción automática

- Evaluación humana de la traducción automática
- Evaluación automática de la traducción automática
- Deficiencias de las medidas de evaluación automática

## 3 Los problemas de la traducción automática

## 4 Funcionamiento de la traducción automática

## 5 Técnicas de traducción automática

- Traducción automática basada en reglas
- Traducción automática basada en corpus

## 1 Traducción automática

- Aplicaciones de la traducción automática
- Automatización del proceso de traducción

## 2 Evaluación de la traducción automática

- Evaluación humana de la traducción automática
- Evaluación automática de la traducción automática
- Deficiencias de las medidas de evaluación automática

## 3 Los problemas de la traducción automática

## 4 Funcionamiento de la traducción automática

## 5 Técnicas de traducción automática

- Traducción automática basada en reglas
- Traducción automática basada en corpus

¿Qué entendemos por **traducción** ?

¿Qué entendemos por **traducción** ?

**Proceso:** Producción de textos en una lengua meta (LM) basado, entre otros, en la modificación de un texto en lengua origen (LO) para hacerlos adecuados a un **propósito**

**Producto:** Documento en LM derivado de otro en LO y que mantiene cierta similitud con éste (contenido, estilo, ...)

Definición de traducción (Sager, 1993):

- rango de actividades humanas **deliberadas**,
- que se hacen como resultado de **instrucciones** recibidas de un **tercero**, y

Definición de traducción (Sager, 1993):

- rango de actividades humanas **deliberadas**,
- que se hacen como resultado de **instrucciones** recibidas de un **tercero**, y
- que consisten en la **producción** de **textos** en una **LM**,
- basada, entre otras cosas, en la **modificación** de un **texto** en una **LO**

Definición de traducción (Sager, 1993):

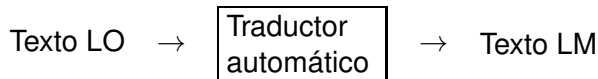
- rango de actividades humanas **deliberadas**,
- que se hacen como resultado de **instrucciones** recibidas de un **tercero**, y
- que consisten en la **producción** de **textos** en una **LM**,
- basada, entre otras cosas, en la **modificación** de un **texto** en una **LO**
- para hacerlo **adecuado** a un **nuevo propósito**.

¿Y por traducción automática?



¿Y por **traducción automática**?

La **traducción**, mediante un **sistema informático** (ordenador + programas) de **textos** en formato digital de una lengua a otra



Sin intervención humana de ningún tipo

- 1 Traducción automática
  - Aplicaciones de la traducción automática
  - Automatización del proceso de traducción
- 2 Evaluación de la traducción automática
  - Evaluación humana de la traducción automática
  - Evaluación automática de la traducción automática
  - Deficiencias de las medidas de evaluación automática
- 3 Los problemas de la traducción automática
- 4 Funcionamiento de la traducción automática
- 5 Técnicas de traducción automática
  - Traducción automática basada en reglas
  - Traducción automática basada en corpus

Una traducción automática, ¿es realmente una traducción?

Una traducción automática, ¿es realmente una traducción?

- No puede ser usada como lo sería una traducción profesional
- ¡Esto no quiere decir que la traducción automática sea *inútil!*

La traducción automática y la traducción profesional, aunque cercanas y relacionadas en su finalidad, no son productos intercambiables (Sager, 1994)

- ¿Qué usos y campos de aplicación puede tener la traducción automática?
- ¿Es necesario que un sistema de traducción automática siempre haga traducciones *perfectas* de cualquier texto para ser útil?

irlandés→inglés

Is ar an Oileán Fada a bhí mé féin agus m'fhear céile ag fanacht nuair a rinne muid an turas sin. Dúradh linn fanacht cois farraige ag am díthrá agus thuirling an muireitleán anuas chun sinn a thabhairt ar bord.

On the Long Island I and my husband waiting when we made the trip. We were told seaside stay at low tide and landed the Flying down to us on board.

irlandés→inglés

Is ar an Oileán Fada a bhí mé féin agus m'fhear céile ag fanacht nuair a rinne muid an turas sin. Dúradh linn fanacht cois farraige ag am díthrá agus thuirling an muireitleán anuas chun sinn a thabhairt ar bord.

On the Long Island I and my husband waiting when we made the trip. We were told seaside stay at low tide and landed the Flying down to us on board.

## Asimilación

Traducción **efímera**, idealmente instantánea, para la revisión o la comprensión de documentos escritos en una lengua que no se domina

- navegación por Internet, comunicación informal, etc.

francés→inglés

Au cours de ses 43 années d'aventure européenne, Londres a souvent été perçu comme réticent à tout nouvel approfondissement de l'Union européenne et à une intégration plus avancée. Volontairement en dehors de la zone euro et de l'espace Schengen, le pays a régulièrement critiqué les institutions européennes et remis en cause sa contribution au budget communautaire.

During its 43 years of European Adventure, London has often been seen as reluctant to any further deepening of the European Union and further integration. Voluntarily outside the euro area and Schengen space, the country has regularly criticized the European institutions and undermined its contribution to the EU budget.



francés→inglés

Au cours de ses 43 années d'aventure européenne, Londres a souvent été perçu comme réticent à tout nouvel approfondissement de l'Union européenne et à une intégration plus avancée. Volontairement en dehors de la zone euro et de l'espace Schengen, le pays a régulièrement critiqué les institutions européennes et remis en cause sa contribution au budget communautaire.

During its 43 years of European Adventure, London has often been seen as reluctant to any further deepening of the European Union and further integration. Voluntarily outside the euro area and Schengen space, the country has regularly criticized the European institutions and undermined its contribution to the EU budget.

## Diseminación (postedición)

Traducción **permanente**, idealmente con pocos errores, para su publicación después de su **corrección**

- Producción de borradores para su **postedición**

¿Hay alguna diferencia?

	Necesario	Innecesario
Diseminación	Sintaxis precisa Errores predecibles Precisión alta (WER $\leq$ 20 %) ☺ Traductores felices	Comprensibilidad Traducción rápida
Asimilación	Comprensibilidad Traducción rápida	<i>Corrección</i> sintáctica <i>Corrección</i> léxica Errores predecibles ☺ Traductores felices

Ejemplos de traducción y de su uso:

## Asimilación

Inglés (TA):      *\*Match very difficult but fans unconditional support players very motivated*

---

Ejemplos de traducción y de su uso:

## Asimilación

Inglés (TA): *\*Match very difficult but fans unconditional support players very motivated*

---

Inglés (Cor.): *~~Match~~The game was very difficult but fans the unconditional support of fans made the players to be very motivated*

---

Español (LO): *El partido ha sido muy difícil pero el apoyo incondicional de la afición hizo que los jugadores estuvieran muy motivados*

Ejemplos de traducción y de su uso:

## Diseminación

Inglés (TA):      \**Eat you were not coming we left*

---

Ejemplos de traducción y de su uso:

## Diseminación)

Inglés (TA):      \**Eat you were not coming we left*

---

Inglés (Cor.):    *~~Eat~~ As you were not coming we left*

---

Español (LO):   *Como no venías nos fuimos*

## 1 Traducción automática

- Aplicaciones de la traducción automática
- **Automatización del proceso de traducción**

## 2 Evaluación de la traducción automática

- Evaluación humana de la traducción automática
- Evaluación automática de la traducción automática
- Deficiencias de las medidas de evaluación automática

## 3 Los problemas de la traducción automática

## 4 Funcionamiento de la traducción automática

## 5 Técnicas de traducción automática

- Traducción automática basada en reglas
- Traducción automática basada en corpus

- ¿Cómo podríamos reducir el tiempo necesario para corregir una traducción automática para su disseminación?



- ¿Cómo podríamos reducir el tiempo necesario para corregir una traducción automática para su disseminación?
- ¿Se os ocurre alguna manera de hacer que el sistema cometa menos errores?

## Postedición

Corrección del texto traducido en sucio para hacerlo adecuado al propósito previsto

## Postedición

Corrección del texto traducido en sucio para hacerlo adecuado al propósito previsto

La postedición es conveniente cuando

$$\text{coste(TA)} + \text{coste(postedit)} < \text{coste(trad.manual)}$$

Para que sea eficiente:

- se debe ser competente en la lengua meta → generar un texto genuino a partir del texto en sucio
- se debe conocer el sistema de traducción automática → reconocer el origen de los errores, predecir el comportamiento

## Preedición

Preparación del texto para evitar léxico o construcciones que dan problemas de traducción con un sistema de traducción automática

## Preedición

Preparación del texto para evitar léxico o construcciones que dan problemas de traducción con un sistema de traducción automática

## Algunas reglas

- *Phrasal verbs*: Mantén siempre juntos el verbo y la preposición (o preposiciones)
- Reduce el uso de la pasiva en la medida de lo posible
- Evita el uso de oraciones subordinadas
- Procura usar frases cortas
- Cada frase debe ser sintáctica y semánticamente completa

## Preedición

Preparación del texto para evitar léxico o construcciones que dan problemas de traducción con un sistema de traducción automática

## Algunas reglas

- *Phrasal verbs*: Mantén siempre juntos el verbo y la preposición (o preposiciones)
- Reduce el uso de la pasiva en la medida de lo posible
- Evita el uso de oraciones subordinadas
- Procura usar frases cortas
- Cada frase debe ser sintáctica y semánticamente completa

Conocer bien *cómo funciona* el sistema de traducción automática también ayuda a la preedición

## Lenguaje controlado

Subconjuntos de lenguaje natural resultado de restringir la gramática y el vocabulario para reducir o eliminar la ambigüedad y la complejidad.

### Alternativa a la preedición

- Los autores escriben ya pensando en el tratamiento automatizado del texto
- Se evitan léxico y construcciones problemáticas
- Se minimiza la postedición
- Consistencia de estilo, comprensibilidad, mantenibilidad
- Pero ... ¡los autores lo deben conocer y aplicar!
- Se les puede ayudar con herramientas informáticas

**Texto a traducir:** Computational linguistics is a research area within Artificial Intelligence whose objective is to automatically model humans' linguistic communication ability, that is, the language through which we communicate with each other. It is therefore an interdisciplinary field, in which linguists, computer scientists, and mathematicians, but also experts in a particular domain, such as doctors, psychologists, lawyers, etc. are needed to accurately model the specific characteristics of a specialized context.

## Traductores a probar:

- Apertium: <https://apertium.org/index.cat.html>
- Google Translate: <https://translate.google.com/>
- Bing Translator: <https://www.bing.com/translator>



- ¿Qué traductor os parece mejor? ¿Y peor?

# Ejercicio 1: TA general /2

- ¿Qué traductor os parece mejor? ¿Y peor?
- ¿Qué entendemos por mejor o peor?

# Ejercicio 1: TA general /2

- ¿Qué traductor os parece mejor? ¿Y peor?
- ¿Qué entendemos por mejor o peor?
- ¿Usaríais alguno de estos sistemas para diseminación?

## Ejercicio 1: TA general /2

- ¿Qué traductor os parece mejor? ¿Y peor?
- ¿Qué entendemos por mejor o peor?
- ¿Usaríais alguno de estos sistemas para diseminación?
- ¿Y para asimilación?

- ¿Se os ocurren modificaciones sobre el texto que hagan que Apertium funcione mejor?

## Ejercicio 1: TA general /3

- ¿Se os ocurren modificaciones sobre el texto que hagan que Apertium funcione mejor?
- ¿Qué pasa si aplicáis estos mismos cambios con los otros dos traductores?

- ¿Se os ocurren modificaciones sobre el texto que hagan que Apertium funcione mejor?
- ¿Qué pasa si aplicáis estos mismos cambios con los otros dos traductores?
- ¿Pensáis que vale la pena pre-editar el texto con Apertium?

**Texto a traducir:** Chiquitita, tell me what's wrong  
You're enchain'd by your own sorrow  
In your eyes there is no hope for tomorrow  
How I hate to see you like this  
There is no way you can deny it

### **Traductores a probar:**

- Apertium: <https://apertium.org/index.cat.html>
- Google Translate: <https://translate.google.com/>
- Bing Translator: <https://www.bing.com/translator>



## Ejercicio 2: TA en contexto artístico /2

- ¿Qué traductor os parece mejor? ¿Y peor?
- ¿Qué entendemos por mejor o peor?

## Ejercicio 2: TA en contexto artístico /3

**Versión en español:** Chiquitita, dime por qué  
Tu dolor hoy te encadena  
En tus ojos hay  
Una sombra de gran pena

<https://www.youtube.com/watch?v=Z4oTvuow9cc>

## Ejercicio 2: TA en contexto artístico /4

- ¿Usaríais alguno de estos sistemas para diseminación?
- ¿Y para asimilación?