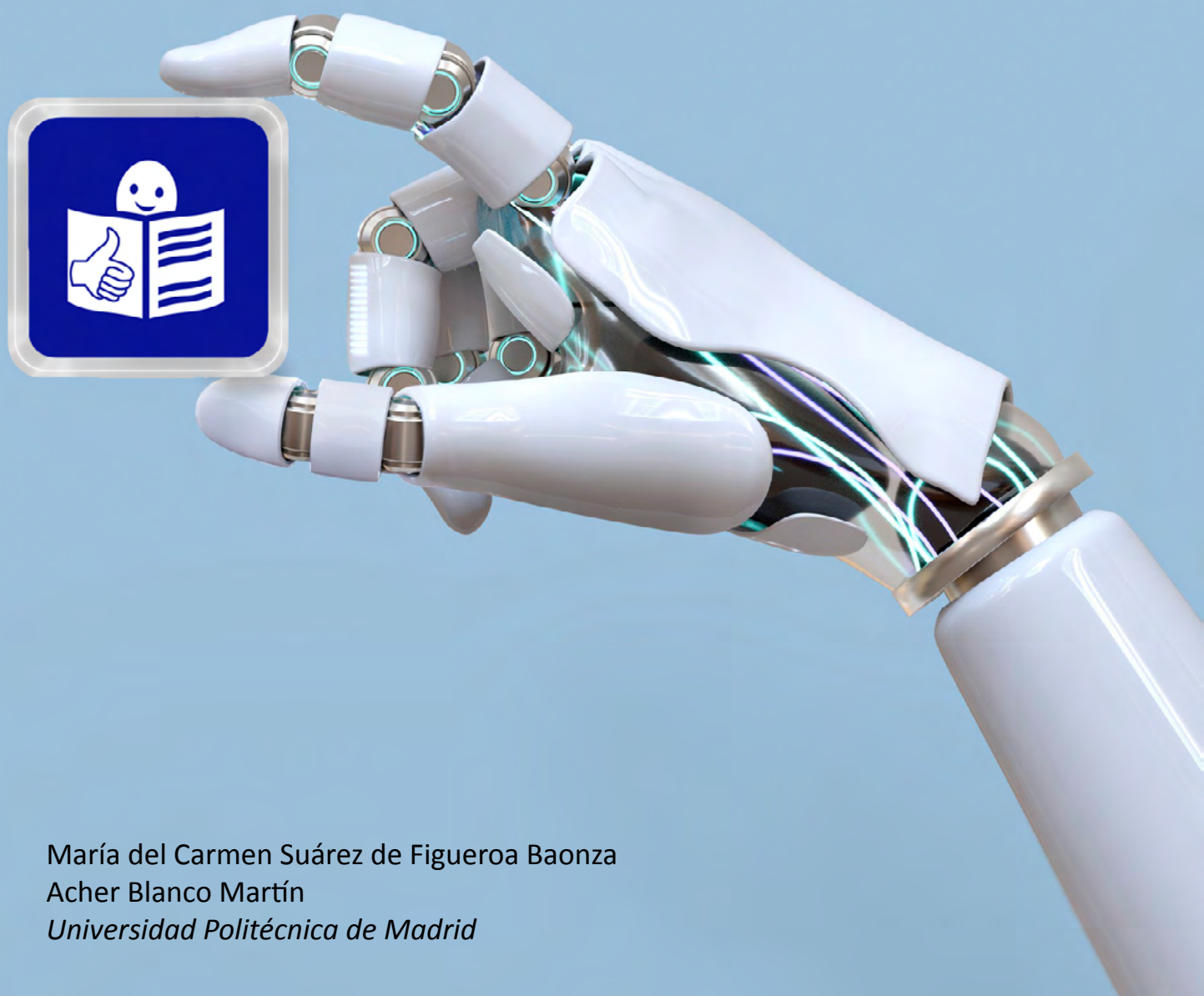


# Informe técnico

## Proyecto de investigación

# Accesibilidad cognitiva e inteligencia artificial

Fase 2



María del Carmen Suárez de Figueroa Baonza  
Acher Blanco Martín  
*Universidad Politécnica de Madrid*

# Informe técnico

## Proyecto de investigación

2023

## Accesibilidad cognitiva e inteligencia artificial. Fase 2



UNIVERSIDAD  
POLITÉCNICA  
DE MADRID





# Índice

## 0. Resumen

## 1. Introducción

## 2. Trabajo Realizado y Resultados

### 2.1 Etapa 1: Establecimiento del protocolo de gestión del piloto

#### 2.1.1 Registro de desarrollo

### 2.2 Etapa 2: Desarrollo de nuevas pautas de lectura fácil

#### 2.2.1 Pauta sobre enumeraciones

#### 2.2.2 Pauta sobre palabras complejas

#### 2.2.3 Pauta sobre palabras indeterminadas

#### 2.2.4 Pauta para evitar sinónimos

#### 2.2.5 Pauta sobre uso del imperativo

#### 2.2.6 Pauta sobre la voz pasiva

### 2.3 Etapa 3: Integración de las nuevas pautas en el sistema piloto

### 2.4 Etapa 4: Difusión del trabajo de investigación

### 2.5 Uso del sistema piloto

#### 2.5.1 Estadísticas de registro

#### 2.5.2 Estadísticas de uso

## 3. Conclusiones

## 4. Líneas Futuras

## 5. Anexo 1: Requisito

### 5.1 Informes de uso

### 5.2 Registro de usuarios

## 6. Anexo 2: Casos de Uso en ATECA

### 6.1 Ejemplos de identificación

### 6.2 Ejemplos de adaptación

## 0. Resumen

El presente documento muestra los resultados obtenidos como fruto de la realización del proyecto de investigación “Accesibilidad Cognitiva e Inteligencia Artificial. Fase II”.

Este proyecto se ha ejecutado en colaboración entre **Plena inclusión**, entidad que gestiona el **Centro Español de Accesibilidad Cognitiva** (CEACOG), y la **Universidad Politécnica de Madrid** (UPM) a través del grupo de investigación **Ontology Engineering Group** (OEG). En este proyecto han participado los siguientes miembros del grupo de investigación Ontology Engineering Group (OEG):

- 
- Dra. María del Carmen Suárez de Figueroa Baonza (Investigadora Principal del Proyecto)
  - D. Isam Diab Lozano (Estudiante de Doctorado)
  - D. Alejandro Muñoz Navarro (Desarrollador de Software)
  - D. Acher Blanco Martín (Desarrollador de Software)
  - Dña. María García-Agudo Estepa (Desarrolladora de Software)
- 

Las actividades del proyecto se han desarrollado en el periodo comprendido entre agosto de 2023 y diciembre de 2023, de conformidad con la cláusula decimoquinta del Convenio suscrito entre el Real Patronato y Plena Inclusión. Durante el año 2024, se garantiza el acceso continuado e ininterrumpido, en días laborables y horario laboral, al piloto resultado del proyecto, sin que esto suponga ningún aporte económico adicional a lo establecido en el convenio.

El documento está organizado en las siguientes secciones: introducción, trabajo realizado y resultados (agrupado por las diferentes etapas del proyecto), conclusiones y líneas futuras.

Además, este documento incluye dos secciones dedicadas a anexos: una relativa a los requisitos establecidos para el registro de usuarios y los informes de estadísticas de uso, y otra referente a ejemplos de casos de uso del piloto software desarrollado en el contexto de este proyecto.

# 1. Introducción

El proyecto de investigación propuesto en este documento tiene **dos objetivos** claramente identificados:

1. **Demostrar** que un sistema de apoyo, basado en Inteligencia Artificial (IA), para realizar el proceso global de adaptación de materiales a las pautas indicadas en la Metodología de Lectura Fácil es de utilidad y facilita el trabajo diario a los adaptadores de documentos y
2. **Analizar** diferentes enfoques para incorporar la automatización de un conjunto de pautas de la metodología de lectura fácil.

En base a los objetivos mencionados, el proyecto de investigación “Accesibilidad Cognitiva e Inteligencia Artificial. Fase II” presenta el desafío de incorporar al proceso tradicional de adaptación manual de documentos a la Metodología de Lectura Fácil una aplicación piloto que realiza de manera (semi)-automática las dos primeras actividades del proceso cíclico de adaptación de documentos a Lectura Fácil<sup>1</sup>: el análisis de documentos y la sugerencia de transformación de dichos documentos para que cumplan con las pautas de dicha metodología.

Este proyecto de investigación se ha desarrollado en **4 etapas** para permitir alcanzar los objetivos planteados.

• **Etapla 1:** Establecimiento del protocolo de gestión del piloto. En esta etapa se han realizado las actividades relacionadas con el protocolo de gestión del piloto desarrollado en Fase I y mejorado en Fase II.

• **Etapla 2:** Desarrollo de nuevas pautas de lectura fácil. En esta etapa se ha trabajado con las siguientes pautas:

- Pautas relativas a la ortotipografía:
  - Se deberían utilizar los dos puntos (:) cuando se introduce una lista que enumera más de tres elementos.
- Pautas relativas al vocabulario y las expresiones:
  - Se debería evitar el uso de palabras muy largas o que contengan sílabas complejas.
  - Se debería evitar el uso de palabras de contenido indeterminado como: cosa, algo o asunto.
  - Se debe utilizar a lo largo del texto la misma palabra para denominar el mismo objeto o referente.
- Pautas relativas a las frases y oraciones:
  - Se debería utilizar el imperativo solo en contextos claros para evitar que se confunda con la tercera persona del singular del presente de indicativo.
  - Se debe evitar la voz pasiva<sup>2</sup>.

1 UNE 153101:2018 EX. Lectura Fácil. Pautas y recomendaciones para la elaboración de documentos.

2 Esta pauta se inició en Fase I y se ha mejorado en Fase II.

• **Etapla 3:** Integración de las nuevas pautas en el sistema piloto. En esta etapa se han integrado las pautas desarrolladas en la Etapa 2 en la aplicación web piloto, denominada “ATECA: Adaptación de Textos en CAstellano” (desarrollada en Fase I y mejorada en esta Fase II).

• **Etapla 4:** Difusión del trabajo de investigación. En esta etapa se han realizado diversas actividades relacionadas con la difusión del trabajo de investigación.

En el resto del documento se presentan:

- (a) un resumen del trabajo realizado y sus resultados,
- (b) unas conclusiones del trabajo realizado en el contexto del proyecto y
- (c) un listado de posibles trabajos futuros en el contexto de la utilización de la Inteligencia Artificial para mejorar la Accesibilidad Cognitiva.

Además, el documento contiene dos anexos:

- (1) uno que incluye documentos de requisitos sobre el sistema de registro y el sistema de recopilación de estadísticas y
- (2) otro que muestra ejemplos de casos de uso de ATECA.



## 2. Trabajo realizado y resultados

### 2.1

#### Etapa 1: Establecimiento del protocolo de gestión del piloto

En esta etapa se ha realizado tanto el diseño del acceso al piloto como el diseño de la analítica web para el piloto.

El equipo de trabajo de la UPM se ha reunido en diversas ocasiones con el equipo de trabajo del CEACOG con el objetivo de establecer los requisitos tanto para la gestión del acceso al piloto como para el diseño de la analítica web.

Una vez se tenían claros los requisitos, se diseñó un primer interfaz de acceso. Dicho interfaz fue validado por el equipo de trabajo del CEACOG.

El sistema de acceso al piloto se basa en un proceso de registro de usuario y un proceso de inicio de sesión. La Figura 1 presenta la página inicial de registro de usuarios en ATECA; mientras que la Figura 2 muestra la página de inicio de sesión.

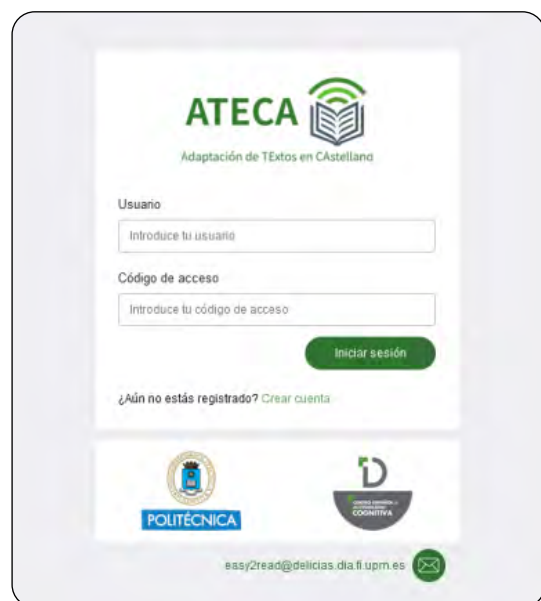
Figura 1. Página de registro en ATECA



Enlace para acceder a ATECA:



Figura 2. Página de inicio de sesión en ATECA



El proceso de registro se lleva a cabo a través de un formulario en el que se deben responder a varias preguntas. Estas han sido creadas para recoger la información relevante de los usuarios, que es la siguiente:

- País de residencia y comunidad autónoma (si aplica)
- Género
- Uso previsto de la aplicación (profesional, personal u otro)
- Experiencia previa en Lectura Fácil
- Pertenencia al Tercer Sector, en cuyo caso:
  - Ámbito de trabajo de la organización
  - Tipos de discapacidad con los que se trabajan (si aplica)
  - Pertenencia a Plena Inclusión (si se trabaja con discapacidad intelectual)
  - Vínculo con la organización (profesional, voluntario, usuario u otro)
- Pertenencia a la administración pública
- Nombre de la organización / administración pública

Adicionalmente, al finalizar este proceso, el usuario revisa y acepta las políticas de protección de datos (disponibles [aquí](#)) y da su consentimiento para tratar los datos recogidos con fines de investigación.

### 2.1.1 Registro de desarrollo

Los requisitos iniciales de los sistemas de acceso y analítica fueron establecidos en los documentos que aparecen en el Anexo 1. Estos requisitos han sido levemente modificados durante el desarrollo del proyecto, a través de diferentes reuniones y mensajes de coordinación. A continuación se presentan los cambios más relevantes:

- **27/09/2023:** se valida la primera versión del formulario de registro y se acuerda reemplazar nombre y apellido por nombre de usuario, para evitar manejar datos identificativos sensibles. También se confirma que la aplicación contará con protocolo HTTPS y se propone al equipo del CEACOG el uso de Google Analytics<sup>3</sup> como sistema de analítica.

<sup>3</sup> [Enlace](#).

- **22/11/2023:** se acuerdan con el equipo del CEACOG las preguntas de la versión final del formulario, su orden y su dependencia (qué preguntas llevan a otras).
- **24/11/2023:** se valida la versión más reciente del registro, se confirma el texto de protección de datos a utilizar y se acuerda añadir una casilla de consentimiento informado para los usuarios.
- **05/12/2023:** visto bueno por parte del equipo del CEACOG para utilizar Google Analytics como sistema de analítica.
- **19/12/2023:** se acepta con el equipo del CEACOG modificar el formulario para recoger dos datos adicionales que no aparecían en el documento inicial sobre el registro, pero sí en el de estadísticas (comunidad autónoma del usuario y nombre de su organización en caso de pertenecer al Tercer Sector).
- **20/12/2023:** se reciben los logos finales que deben aparecer en la aplicación.
- **23/01/2024:** se planifica una reunión de formación con el equipo del CEACOG para enseñar el uso y funcionamiento de Google Analytics como herramienta de análisis y estadísticas. Se comparten las credenciales de acceso a la cuenta ateca.lectura.facil@gmail.com, donde se encuentra dicha configuración.



En esta etapa se ha realizado un análisis de viabilidad con el objetivo de investigar si es factible la automatización de las pautas y recomendaciones de lectura fácil incluidas en este proyecto. Para ello el equipo de trabajo de la UPM ha estudiado de forma exhaustiva las características de cada pauta. En el análisis de viabilidad mencionado se ha seguido una metodología incremental en la que se han desarrollado métodos, basados en IA, para la automatización de:

1. La identificación de las problemáticas mencionadas en las pautas seleccionadas.
2. La propuesta de sugerencias de adaptación de textos en castellano para que las pautas identificadas se cumplan.

Todos los métodos desarrollados se han implementado en Python. En todos los casos los métodos desarrollados se han basado en técnicas de Inteligencia Artificial simbólica<sup>4</sup>, como la representación de conocimiento en forma de reglas y el procesamiento de lenguaje natural.

A continuación, se describe en detalle el trabajo realizado con respecto al desarrollo de las pautas seleccionadas.

### 2.2.1 Pauta sobre enumeraciones

La pauta sobre enumeraciones indica lo siguiente:

“Se deberían utilizar los dos puntos (:) cuando se introduce una lista que enumera más de tres elementos”

Adaptando las indicaciones de la pauta 6 de Ortotipografía (“Se deberían utilizar los dos puntos (:) cuando se introduce una lista que enumera más de tres elementos.”) y la pauta 6 de Texto y Estilo (“Si se van a enumerar los elementos relacionados con una idea en una sola frase, se debería utilizar la coma para separarlos. Siempre que sea posible, si hay varios elementos (como una lista de tareas), se debe utilizar un listado.”) de la Norma UNE de Lectura Fácil, se ha decidido señalar como problemáticas las enumeraciones lineales de 4 o más elementos y adaptarlas como listados con viñetas. Para identificarlas se hace uso de expresiones regulares (que buscan los patrones típicos) y relaciones de dependencia sintáctica, que confirman o desmienten si un patrón detectado es una estructura enumerativa. La adaptación hace uso de información recopilada durante la identificación para colocar los dos puntos en el lugar adecuado, y después segmenta la enumeración, colocando cada elemento en una viñeta del listado.

Uno de los objetivos transversales de este proyecto es conseguir un proyecto de investigación en IA inclusivo. Para ello, se ha incorporado en esta etapa del proyecto la participación de personas con dificultades de comprensión lectora así como la participación de expertos en lectura fácil. Dicha participación se ha hecho efectiva mediante el uso de cuestionarios on-line. Estos cuestionarios se enviaron de forma masiva el jueves 2 de noviembre de 2023, y se estableció como fecha tope para rellenar dichos cuestionarios el viernes 10 de noviembre de 2023.

4 Norvig, P. & Russell, S. J. (2021). Artificial Intelligence: A modern approach (4th Edition). Pearson Global Editions. ISBN: 9781292401133; ISBN ebook: 9781292401171.

Por un lado, en el caso de las personas con dificultades de comprensión lectora (principalmente personas con discapacidad cognitiva), se ha lanzado un cuestionario<sup>5</sup> para obtener información sobre que opción de adaptación de enumeraciones es más fácil y se entiende mejor. Este cuestionario ha sido contestado por 164 personas con discapacidad cognitiva. Los resultados de este proceso han servido para tomar decisiones sobre la manera más adecuada de adaptar enumeraciones.

Por otro lado, en el caso de los expertos en lectura fácil, el cuestionario<sup>6</sup> enviado tenía como objetivo validar las adaptaciones realizadas por nuestra implementación de la pauta. Este cuestionario ha sido contestado por 73 expertos en lectura fácil. Los resultados obtenidos han servido para matizar y mejorar la implementación de esta pauta.

La implementación de esta pauta se ha evaluado con un conjunto de 121 frases extraídas del BOE y de artículos del periódico Público.

## 2.2.2 Pauta sobre palabras complejas

La pauta sobre palabras complejas indica lo siguiente:

---

“Se debería evitar el uso de palabras muy largas o que contengan sílabas complejas”

---

Con el objetivo de entender esta pauta y seguir en la línea de realizar un proyecto de investigación en IA inclusivo, hemos incorporado la participación de expertos en lectura fácil, específicamente adaptadores. Para ello, hemos creado un cuestionario<sup>7</sup> on-line para extraer datos relevantes sobre como los adaptadores realizan su trabajo de adaptación a lectura fácil; en concreto, como realizan la adaptación de palabras complejas.

El mencionado cuestionario se lanzó el miércoles 22 de noviembre de 2023, y se estableció como fecha tope para rellenar dicho cuestionario el viernes 1 de diciembre de 2023.

El cuestionario tiene 3 secciones:

1. Preguntas sobre la longitud de las palabras
2. Preguntas sobre la complejidad de las sílabas
3. Preguntas generales sobre la pauta.

Este cuestionario ha sido contestado por 21 adaptadores de lectura fácil.

Los resultados de dicho cuestionario han permitido tomar decisiones sobre lo que se debe considerar palabra compleja:

- Palabras con 5 o más sílabas (excepto si se encuentran entre las 2000 palabras más frecuentes<sup>8</sup>)
- Palabras con 3 consonantes seguidas (excepto si se encuentran entre las 2000 palabras más frecuentes)
- Palabras con 4 consonantes seguidas (excepto si se encuentran entre las 1000 palabras más frecuentes)
- Palabras con patrones de sílaba CCVCC<sup>9</sup>, CVCC y VCC (excepto si se encuentran entre las 2000 palabras más frecuentes)

---

5 [Enlace](#)

6 [Enlace](#)

7 [Enlace](#)

8 “Lista total de frecuencias” [aquí](#)

9 C significa constante y V vocal.

Siguiendo estos criterios se pueden identificar palabras potencialmente problemáticas para la accesibilidad cognitiva. Una vez detectadas, la adaptación realizada consiste en sustituirlas por un sinónimo que no sea complejo. Esta sustitución se realiza recabando una lista de sinónimos posibles (mediante tesauros y otros recursos), y seleccionando el más adecuado mediante modelos de similitud semántica.

La implementación de esta pauta se ha evaluado con un conjunto de 50 frases, generadas automáticamente con Inteligencia Artificial (ChatGPT).

### 2.2.3 Pauta sobre palabras indeterminadas

La pauta sobre palabras indeterminadas indica lo siguiente:

---

“Se debería evitar el uso de palabras de contenido indeterminado como: cosa, algo o asunto”

---

Esta pauta se centra en la identificación del siguiente patrón en textos en castellano “Verbo ‘ser’ (en cualquier forma verbal) + Palabra ‘cosa’, ‘algo’, ‘asunto’, ‘cosas’ o ‘asuntos’”. Si dicho patrón se detecta en el texto, entonces la pauta no se está cumpliendo. Si se ha detectado alguna palabra de contenido indeterminado en la fase de identificación, dicha palabra se sustituye por un sinónimo adecuado al contexto en fase de sugerencia de adaptación.

Esta pauta utiliza el recurso externo BabelNet para obtener los sinónimos en un determinado contexto. La implementación de esta pauta se ha evaluado con un conjunto de 20 frases creadas por el equipo de pruebas.

### 2.2.4 Pauta para evitar sinónimos

La pauta para evitar sinónimos indica lo siguiente:

---

“Se debe utilizar a lo largo del texto la misma palabra para denominar el mismo objeto o referente”

---

Esta pauta utiliza el mismo sistema de extracción de sinónimos que la pauta de palabras complejas, pero los usa para buscar ocurrencias a lo largo del texto. En caso de que encuentre varias palabras sinónimas entre sí, las marca como problemáticas. La adaptación coge el sinónimo detectado que antes aparece en el texto, y reemplaza los sinónimos posteriores con esa palabra. Por motivos de velocidad y desempeño, solo se analizan sustantivos y nombres propios.

La implementación de esta pauta se ha evaluado con un conjunto de 20 frases, generadas automáticamente con Inteligencia Artificial (ChatGPT).

### 2.2.5 Pauta sobre uso del imperativo

La pauta sobre el uso del imperativo indica lo siguiente:

---

“Se debería utilizar el imperativo solo en contextos claros para evitar que se confunda con la tercera persona del singular del presente de indicativo”

---

La identificación de verbos en imperativo se ha realizado utilizando etiquetas de spaCy<sup>10</sup>. La adaptación se realiza utilizando el siguiente patrón ““Tener” + “que” + “verbo infinitivo”.

La implementación de esta pauta se ha evaluado con un conjunto de 30 frases creadas por el equipo de pruebas.

### 2.2.6 Pauta sobre la voz pasiva

La pauta sobre la voz pasiva indica lo siguiente:

---

“Se debe evitar la voz pasiva”<sup>11</sup>

---

Esta pauta ha sido identificada empleando patrones de etiquetas de partes del discurso (POS) y relaciones de dependencia entre palabras mediante la librería spaCy. Para refinar la precisión de nuestra solución, recurrimos al uso de un modelo basado en transformers (modelo trf). Asimismo, hemos categorizado las identificaciones en dos clases: pasivas con agente y pasivas sin agente.

En lo que respecta a la adaptación, hemos implementado dos enfoques distintos. En el caso de las construcciones pasivas con agente, hemos modificado la estructura de la oración y el verbo para convertirlas a voz activa. Para las construcciones pasivas sin agente, proponemos dos alternativas dada su complejidad:

1. Utilizar la voz pasiva impersonal con “se”, ejemplificado en la transformación de “El concierto fue cancelado” a “Se canceló el concierto”
2. Emplear un sujeto indefinido como “alguien”, transformando “El concierto fue cancelado” a “Alguien canceló el concierto”

Para la evaluación de la pauta, hemos hecho uso de 70 frases redactadas por nosotros o generadas automáticamente con Inteligencia Artificial (ChatGPT).

---

<sup>10</sup> [Enlace](#)

<sup>11</sup> Esta pauta se inició en Fase I y se ha mejorado en Fase II.

En esta etapa se han integrado (a) las pautas desarrolladas en la Etapa 2 del presente proyecto con (b) la interfaz de usuario implementada en la Etapa 1 del proyecto anterior (“Accesibilidad Cognitiva e Inteligencia Artificial”)<sup>12</sup>.

Como resultado de dicha integración se ha obtenido la aplicación web piloto, denominada “ATECA: Adaptación de TExtos en CASTellano”. La arquitectura de la aplicación web piloto se basa en el uso de servicios web.

La Figura 3 muestra la captura de pantalla de ATECA cuando el usuario está usando la funcionalidad de identificar problemáticas con respecto a la Norma UNE de Lectura Fácil; mientras que la Figura 4 muestra el interfaz de la funcionalidad de proponer sugerencias de adaptación.

Además, como parte de esta etapa, se ha realizado el despliegue de ATECA en un servidor web de la UPM. Dicho despliegue se ha producido el día 22 de diciembre de 2023. A partir de ese momento, ATECA está disponible para que los adaptadores de lectura fácil o interesados en la accesibilidad cognitiva puedan utilizar el piloto.

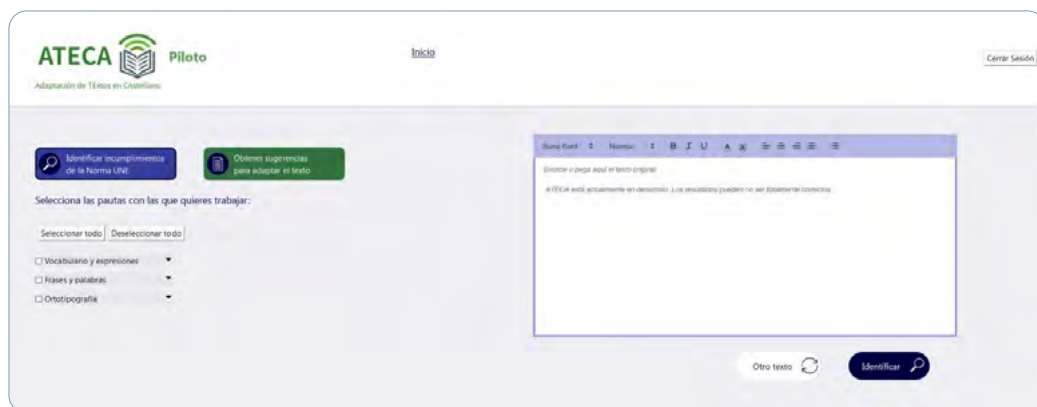


Figura 3. ATECA: Identificación de incumplimientos de la Norma UNE

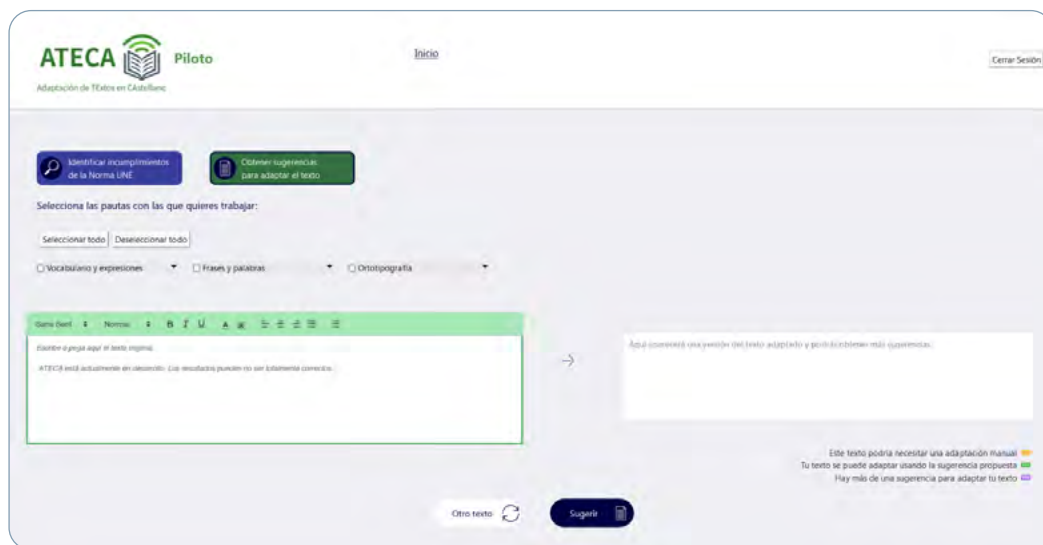


Figura 4. ATECA: Sugerencias de adaptación

<sup>12</sup> Este proyecto se desarrollo entre enero y marzo de 2023.

En esta etapa se han realizado diversas actividades relacionadas con la difusión del trabajo de investigación. Dichas actividades se han realizado en estrecha colaboración entre el equipo de trabajo de la UPM y el equipo de trabajo del CEACOG.

En concreto, se ha impartido una charla de concienciación a profesionales de la lectura fácil para explicar la investigación realizada y las ventajas que se pueden obtener del uso del sistema piloto en el proceso global de adaptación a lectura fácil.

Los objetivos de dicha charla eran:

1. Explicar la investigación y el desarrollo que se está llevando a cabo en la UPM para construir un sistema piloto de ayuda en el proceso de adaptación a lectura fácil
2. Mencionar los beneficios de este sistema piloto, mostrando el alcance actual

Dicha charla tuvo lugar de forma remota el día 20 de noviembre de 2023 utilizando la plataforma Zoom. La charla duró 70 minutos y tuvo un total de 45 asistentes (sin contar a la ponente<sup>13</sup>).

Además, se ha creado una sección de difusión<sup>14</sup> en la página web del CEACOG. En dicha sección se han incluido (a) la explicación de la investigación realizada y (b) el enlace a ATECA, para que pueda usarse por profesionales del ámbito de la discapacidad.

También se ha elaborado una nota de prensa<sup>15</sup> para explicar en detalle el proyecto de investigación y su principal resultado, ATECA.

<sup>13</sup> Mari Carmen Suárez de Figueroa Baonza, como responsable del proyecto de investigación, fue la encargada de presentar el trabajo y el desarrollo del piloto.

<sup>14</sup> [Enlace](#)

<sup>15</sup> [Enlace](#)

El sistema piloto fue desplegado en la dirección <https://ateca.linkeddata.es> el 22/12/2023 para el uso por parte de personas de diferentes perfiles. Adicionalmente, el sistema ha sido conectado a la plataforma Google Analytics<sup>16</sup>, que permite recoger estadísticas de uso y de registro. Gracias a esto se ha generado el siguiente informe con los datos recogidos entre el 8 de enero y el 29 de febrero de 2024.

### 2.5.1 Estadísticas de registro

Gracias a los datos recabados en el cuestionario de registro, podemos conocer diferentes datos sobre los usuarios registrados. Entre las fechas establecidas ha habido un total de 210 registros, que presentan los siguientes perfiles:

- País de residencia: España (92%), Argentina (2%), Colombia (1,5%) y Mexico (1,5%)<sup>17</sup>.
- Comunidades Autónomas (cuando aplica): Madrid (19%), Andalucía (18%), País Vasco (10%), Valencia (9%), Castilla y León (8%), Cataluña (7%) y Baleares (6%)<sup>18</sup>.
- Uso de ATECA: profesional (73%), personal (19%) y otro (9%).
- Experiencia con Lectura Fácil: 49% tiene experiencia previa frente al 51% que no.
- Miembros de Plena Inclusión: 38% pertenece a Plena frente al 62% que no.

---

<sup>16</sup> [Enlace](#)

<sup>17</sup> También presentes Brasil, Chile, Costa Rica, El Salvador, EEUU y Suiza, con menos del 1%

<sup>18</sup> También presentes Aragón, Asturias, La Rioja, Canarias, Galicia, Castilla la Mancha, Navarra, Cantabria y Ceuta con menos del 5%

## 2.5.2 Estadísticas de uso

El uso se encuentra dividido en dos operaciones: identificación y adaptación. Se ha registrado el número total de llamadas a cada operación, así como el número de veces que las pautas disponibles han sido seleccionadas en una operación. Estos datos pueden ser consultados en la Tabla 1. Cabe destacar que una operación puede (y suele) tener seleccionadas más de una pauta simultánea.

| Usos                          | Identificación | Adaptación  |
|-------------------------------|----------------|-------------|
| Pauta Conectores              | 376            | 645         |
| Pauta Enumeraciones           | 378            | 605         |
| Pauta Evitar sinónimos        | 315            | 534         |
| Pauta Uso del imperativo      | 353            | 581         |
| Pauta Mayúsculas              | 356            | 559         |
| Pauta Palabras complejas      | 376            | 560         |
| Pauta Palabras indeterminadas | 372            | 564         |
| Pauta Superlativos            | 414            | 587         |
| Pauta Voz pasiva              | 407            | 678         |
| <b>Totales</b>                | <b>638</b>     | <b>1052</b> |

*Tabla 1. Estadísticas de uso*

La Tabla 1 recoge la actividad de 151 usuarios, lo que representa el 72% de los registros. De los usuarios que sí han utilizado el servicio, se ha recogido una mediana de uso (es decir, la cantidad de operaciones realizadas por usuario más común) de 2 en identificación, y de 3 en adaptación. También sabemos que, de estos 151 usuarios, el 85% ha realizado menos de 10 operaciones totales y el 58% ha realizado menos de 5 operaciones totales.

Estos datos demuestran que, aunque la mayoría de usuarios ha probado el sistema, pocos lo están usando de forma habitual. Sin embargo, tenemos intención de poner en marcha algún sistema que incentive el uso de la aplicación, para poder recabar más datos y recibir comentarios de los usuarios.

### 3. Conclusiones

Como resultado de la realización del proyecto de investigación “Accesibilidad Cognitiva e Inteligencia Artificial. Fase II” se han obtenido, en términos numéricos, los siguientes resultados:

■ **Implementación** de 6 pautas de lectura fácil relacionadas con la redacción de textos. En concreto, las siguientes:

- Se deberían utilizar los dos puntos (:) cuando se introduce una lista que enumera más de tres elementos.
- Se debería evitar el uso de palabras muy largas o que contengan sílabas complejas.
- Se debería evitar el uso de palabras de contenido indeterminado como: cosa, algo o asunto.
- Se debe utilizar a lo largo del texto la misma palabra para denominar el mismo objeto o referente.
- Se debería utilizar el imperativo solo en contextos claros para evitar que se confunda con la tercera persona del singular del presente de indicativo.
- Se debe evitar la voz pasiva.

■ **Desarrollo y despliegue** de una aplicación web piloto llamada ATECA<sup>19</sup> (un sistema de apoyo basado en IA), para realizar el proceso global de adaptación de materiales a las pautas indicadas en la Metodología de Lectura Fácil.

El primer despliegue de ATECA en un servidor de la UPM se realizó el 22 de diciembre de 2024.

La primera versión desplegada de ATECA tenía una serie de limitaciones, que fueron comunicadas a los usuarios potenciales. En concreto:

- Se ha detectado que la pauta de enumeraciones podría tener un funcionamiento extraño al combinarse con otras pautas.
- Se ha descubierto que la pauta de palabras complejas podría devolver adaptaciones no adecuadas, debido al funcionamiento de los recursos externos que se utilizan en dicha pauta.
- Como se ha comentado en la Sección 2.2.3, la pauta de palabras indeterminadas solo identifica un patrón lingüístico concreto<sup>20</sup>. Además, se ha detectado que la sugerencia de adaptación podría resultar no demasiado conveniente en algunos casos, debido a los resultados proporcionados por el recurso externo utilizado.
- Se ha detectado que la pauta para evitar sinónimos presenta un alto número de falsos positivos en la funcionalidad de sugerir adaptaciones. En esos casos, las sugerencias de adaptación podrían provocar confusiones.

Durante los dos primeros meses del año 2024, el equipo de trabajo de la UPM ha analizado las problemáticas mencionadas anteriormente y ha trabajado en la mejora de dichas situaciones. Como resultado de este trabajo, se ha realizado un segundo despliegue de ATECA el 8 de febrero de 2024; siendo esta versión la actual (a fecha de escritura del presente informe).

<sup>19</sup> [Enlace](#)

<sup>20</sup> “Verbo ‘ser’ (en cualquier forma verbal) + Palabra ‘cosa’, ‘algo’, ‘asunto’, ‘cosas’ o ‘asuntos’”

Dicha versión incluye las siguientes **mejoras**:

- Se han solucionado los problemas de colisión de la pauta de enumeraciones con el resto de pautas.
- Se han corregido algunos de los problemas detectados en la pauta de palabras de contenido indeterminado.
- Se han mejorado las sugerencias de adaptación en la pauta sobre el uso del imperativo.
- Se han realizado las siguientes actualizaciones en el interfaz de usuario de ATECA:
  - Nuevo botón para deseleccionar las pautas seleccionadas.
  - Nuevo color para indicar una problemática en el texto para la que ATECA no ha podido sugerir una adaptación adecuada.

No obstante, en la versión de ATECA desplegada el 8 de febrero de 2024 no se han podido solventar una serie de errores disruptivos en dos de las pautas implementadas: evitar sinónimos y palabras complejas. En dicha versión de ATECA, estas pautas fallan en su funcionamiento, y debido a la configuración del sistema, detienen la ejecución, aunque haya otras pautas seleccionadas que sí funcionan. En este sentido, el equipo de trabajo de la UPM está desarrollando una nueva versión de ATECA que soluciona estas problemáticas como se puede observar en las Figuras 14, 16, 24, 25 y 26 del Anexo 2. Esta nueva versión, que se desplegará en las próximas semanas, también solucionará algunos casos de error conocidos, mejorará el funcionamiento de diversas pautas y modificará la configuración del sistema para que el fallo de una pauta no afecte a las demás pautas seleccionadas.

Es importante destacar que debido a que el despliegue de ATECA se tuvo que realizar en la penúltima semana del mes de diciembre de 2023, la evaluación del sistema piloto no se ha podido realizar todavía.

Sin embargo, el equipo de trabajo de la UPM está esbozando un mapa de actuación con respecto al lanzamiento de diferentes cuestionarios para recabar opiniones sobre el uso y la utilidad de ATECA. Los cuestionarios, principalmente on-line, se dirigirán principalmente a adaptadores pertenecientes al CEACOG y a Plena Inclusión, responsables del CEACOG impulsores del proyecto de colaboración, así como agentes externos interesados por la accesibilidad cognitiva.

Finalmente, mencionar, que el equipo de trabajo de la UPM ha trabajado (a) en la escritura de un artículo de divulgación científica, que será enviado a la revista científica “Siglo Cero”<sup>21</sup> (promovida por INICO - Universidad de Salamanca y Plena inclusión) en los próximos meses; y (b) en la planificación de la escritura de artículos de investigación para ser enviados a conferencias o revistas científicas en el segundo trimestre de 2024.

---

<sup>21</sup> [Enlace](#)

## 4. Líneas futuras

Uno de los resultados tangibles del proyecto de investigación “Accesibilidad Cognitiva e Inteligencia Artificial. Fase II” es la aplicación web piloto ATECA, que se puede considerar un sistema de ayuda a la decisión para facilitar el proceso de adaptación de documentos siguiendo la Metodología de Lectura Fácil.

Durante el año 2023, se han realizado diferentes tareas que han dado como resultado el piloto ATECA, que se está usando actualmente por diferentes perfiles de usuario.

Con el objetivo de mejorar algunos aspectos que no pudieron tratarse completamente durante el desarrollo del proyecto, se plantean las siguientes actividades como trabajos futuros **a corto plazo**:

- **Ciclo de seminarios sobre ATECA.** Esta actividad tiene dos objetivos: (1) explicar en detalle el funcionamiento del piloto y sus limitaciones actuales; (2) mostrar diferentes casos de uso de ATECA; y (3) resolver dudas o comentarios por parte de los usuarios.
- **Evaluación de la usabilidad y la utilidad de ATECA.** El objetivo de esta actividad es recopilar opiniones de los usuarios sobre como de usable y como de útil es actualmente ATECA.
- **Evaluación de las funcionalidades de ATECA,** así como de cada pauta de manera independiente. El objetivo de esta actividad es mejorar el piloto desde una perspectiva global así como desde una perspectiva local.
- **Actividades de difusión de ATECA.** En este caso se pretende (a) establecer un plan de lanzamiento de retos durante el año 2024 para motivar a los adaptadores a usar ATECA y a proporcionar sus impresiones y comentarios; (b) elaborar un documento sencillo que describa las pautas incluidas en ATECA así como los pasos básicos para el uso de dicho piloto; y (c) planificar la escritura de artículos de investigación, artículos de divulgación científica, y notas de prensa, así como la participación en eventos de investigación y divulgación.

Por otro lado, se plantean las siguientes líneas futuras **a medio y largo plazo**:

- **Mejora continua de ATECA** en cuanto a los siguientes aspectos: (a) actualización de las pautas incluidas actualmente en el piloto y (b) inclusión del tratamiento automatizado de nuevas pautas de la metodología de lectura fácil.
- **Análisis de la necesidad de tener aplicaciones específicas,** como ATECA, en base al dominio (médico, legal, educativo, por mencionar algunos ejemplos) y/o al tipo de documento (literario, divulgativo, entre otros).
- **Creación de protocolos y metodologías** para garantizar que los proyectos de investigación basados en Inteligencia Artificial sean inclusivos.  
Uno de los puntos cruciales para conseguir dicha inclusividad consiste en incorporar a personas con dificultades de comprensión lectora en todas las actividades del proyecto, tanto en talleres de co-creación como en pruebas de uso y accesibilidad.

Este listado de trabajos futuros, en la mayoría de los casos, se basa en la realización de una investigación multidisciplinar que aúne lingüística, cognición y tecnología y que involucre a profesionales de distintas áreas (ingenieros informáticos, lingüistas computacionales, expertos en lectura fácil, entre otros).

## 5. Anexo 1: Requisitos

Los siguientes requisitos han sido extraídos de los documentos iniciales recibidos por parte del equipo del CEACOG durante la coordinación inicial del proyecto:

Requisitos-InformeUsos y Requisitos-RegistroUsuarios.

### 5.1

#### Informes de uso

##### 3.- INFORMACIÓN REPORTE

Para explotar la información.

Datos que necesitamos vía consulta o informes que estén programados en la aplicación:

- Fecha de alta de cada persona
- Datos de los **usuarios**:
  - Numero total, numero por perfiles, etc.
  - Listado entidades de donde son los usuarios
  - Numero de personas de entidades de Plena y numero de personas de otras entidades
  - Numero de personas que tienen experiencia
  - Por territorios: Comunidades autonómicas y países
- Datos de los **usos** que se desarrollan en el aplicativo:
  - Número y tipología de usos utilizados por los usuarios.
- ¿Se puede saber qué personas se han dado de alta y nunca han hecho nada en el aplicativo?

Figura 5. Documento de requisitos de informes de uso.

**a) Datos que necesitamos de las personas interesadas en usar la aplicación:**

1. Nombre y apellidos
2. ¿Para qué quieres usar la aplicación?
  - Para uso personal.
  - Para uso profesional.
  - Para otro uso.
3. Sexo
  - Mujer
  - Hombre
  - Otro (rellenar)
  - Prefiero no decirlo
4. País de residencia (desplegable).
5. ¿Tienes experiencia en la adaptación de textos en Lectura Fácil?
  - No
  - Sí
6. ¿Formas parte de alguna organización del Tercer Sector)
  - No (pasa a pregunta 6)
  - Sí (pasa a pregunta 7)
7. Ámbito de trabajo de tu organización:  
(Desplegable)
  - Discapacidad (desplegable con tipo de discapacidad y, en caso de que diga discapacidad intelectual y del desarrollo:
    - ¿Forma parte de Plena inclusión? Sí/No (en caso afirmativo, desplegable).
  - Resto de ámbitos
8. Vínculo con tu organización (desplegable)
  - Profesional.
  - Persona voluntaria.
  - Recibo apoyo de la organización.
  - Otro perfil (rellenar)
9. ¿Trabajas en la Administración Pública?
  - No
  - Sí (nuevo campo, indica en qué Administración Pública trabajas).
10. Email con el que quieres crear la cuenta
11. Contraseña de tu cuenta

Casillas de protección de datos

Casilla de envío de información del CEACOG

*Figura 6. Documento de requisitos de registro.*

## 6. Anexo 2: Casos de Uso en ATECA

### 6.1

#### Ejemplos de identificación

The screenshot displays the ATECA web application interface. At the top, the logo "ATECA Piloto" is visible, along with the text "Adaptación de Textos en Castellano". A navigation bar includes "Inicio" and "Cerrar Sesión".

The main content area features two buttons: "Identificar incumplimientos de la Norma UNE" and "Obtener sugerencias para adaptar el texto". Below these, a section titled "Selecciona las pautas con las que quieres trabajar:" contains a "Seleccionar todo" button and a list of checkboxes:

- ☐ Vocabulario y expresiones
  - ☐ Palabras complejas
  - ☒ Superlativos
  - ☐ Palabras indeterminadas
  - ☐ Evitar sinónimos
- ☐ Frases y palabras
- ☐ Ortotipografía
  - ☒ Mayúsculas
  - ☐ Enumeraciones

To the right, a text editor shows a sample text about Santiago Scanlan. Below the editor, a green button reads "Hay un informe a continuación sobre las pautas de lectura fácil".

At the bottom, there are buttons for "Otro texto" and "Identificar". Below these, a red-bordered box highlights the "Superlativos" section, indicating a violation of the UNE norm. It states: "La pauta 8 de la recomendación sobre Vocabulario y expresiones de la UNE indica 'Se deben evitar los superlativos. Es recomendable añadir el adverbio muy al adjetivo o al adverbio.'" and lists the elements that do not follow the norm: "rapidísimo" and "muchísimas posibilidades".

A green-bordered box below highlights the "Mayúsculas" section, indicating a violation of the UNE norm.

The footer contains logos of various institutions, including Ontology Engineering Group, PONTÉCNICA, and U23, along with contact information: "CONTACTO easy2read@delicias.dia.fi.upm.es".

Figura 7. Identificación de superlativos y mayúsculas.

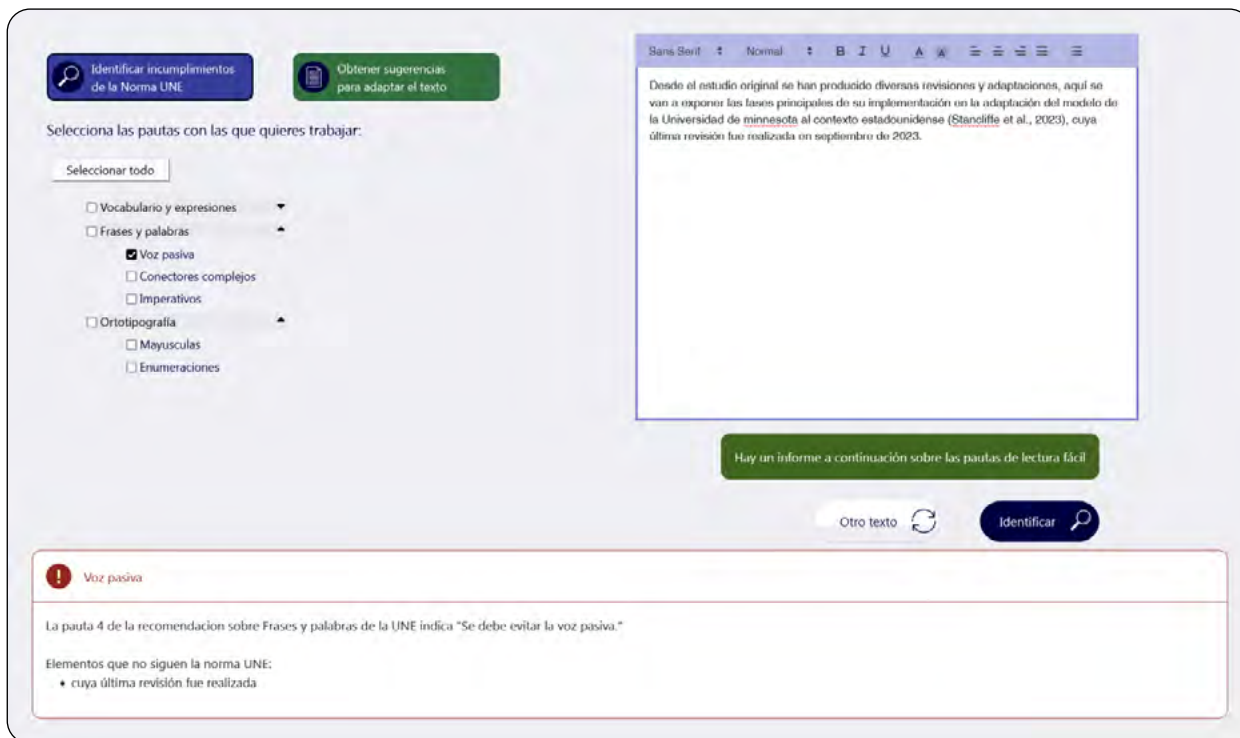


Figura 8. Identificación de voz pasiva.

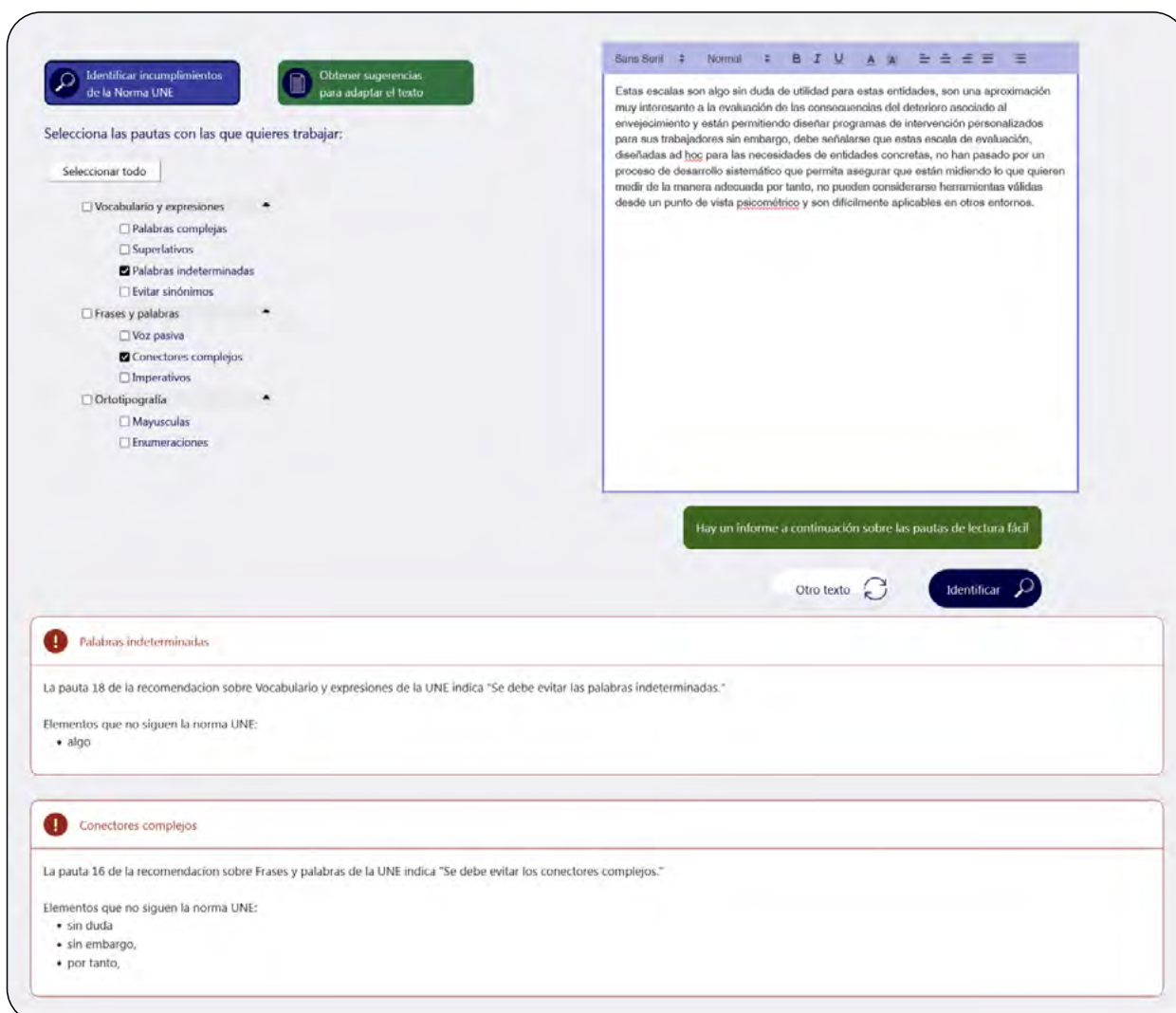


Figura 9. Identificación de palabras indeterminadas y conectores complejos.

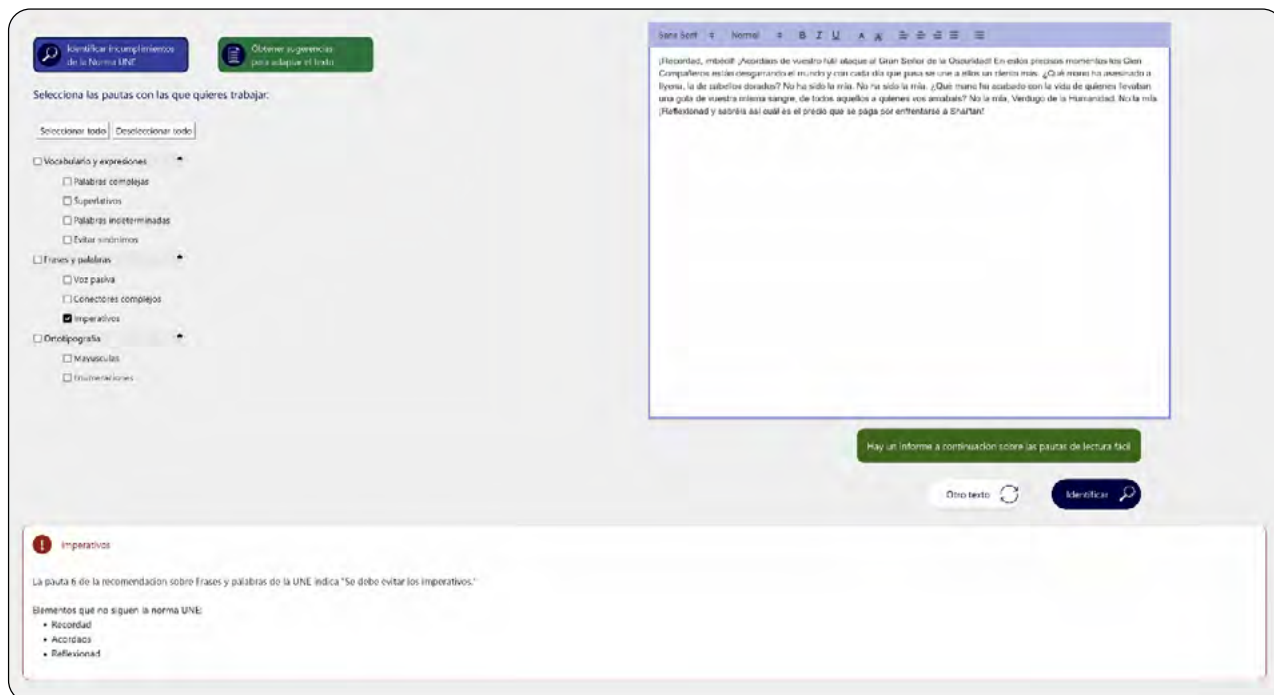


Figura 10. Identificación de imperativos.

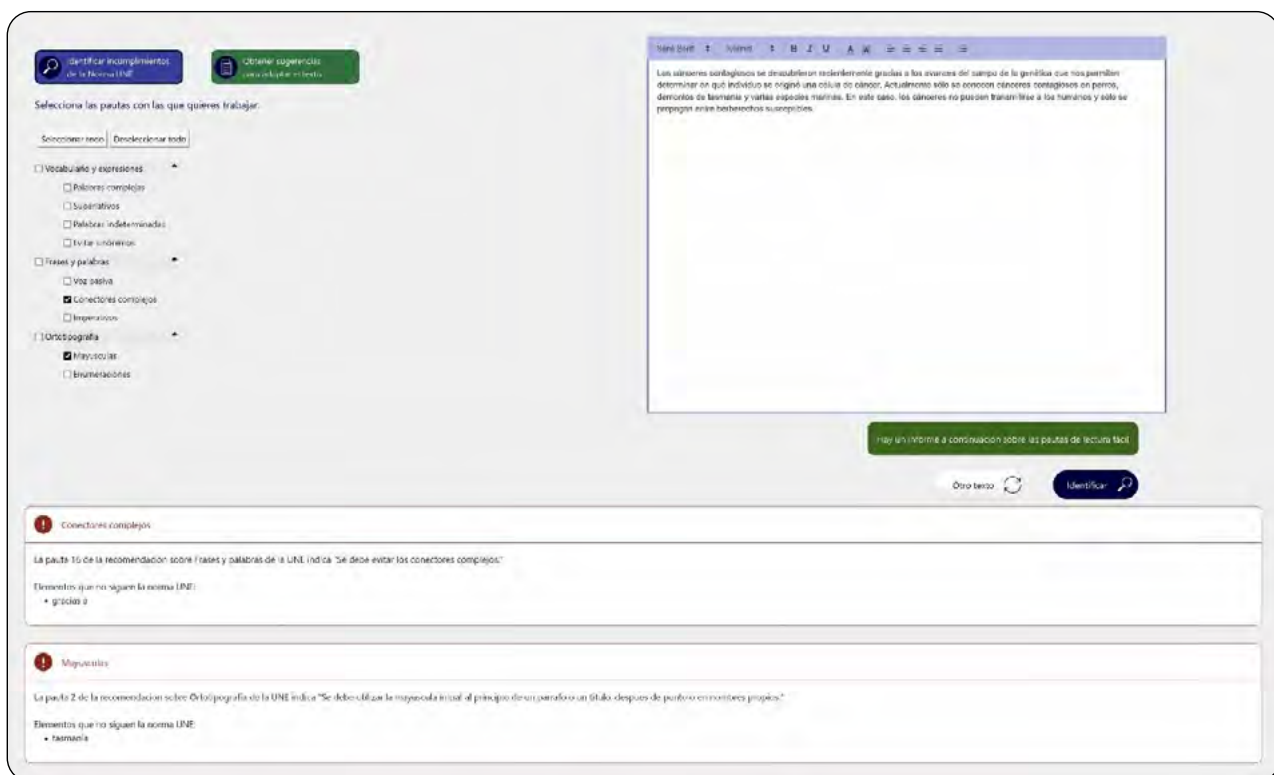


Figura 11. Identificación de conectores complejos y mayúsculas.

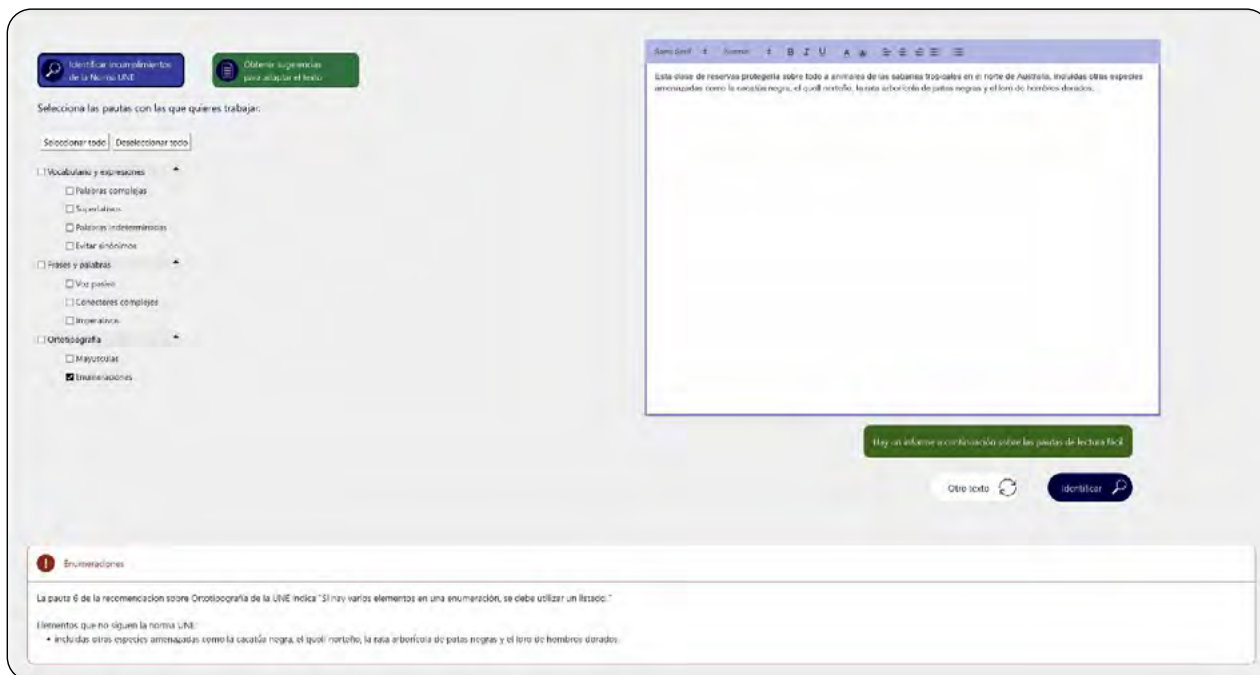


Figura 12. Identificación de enumeraciones.

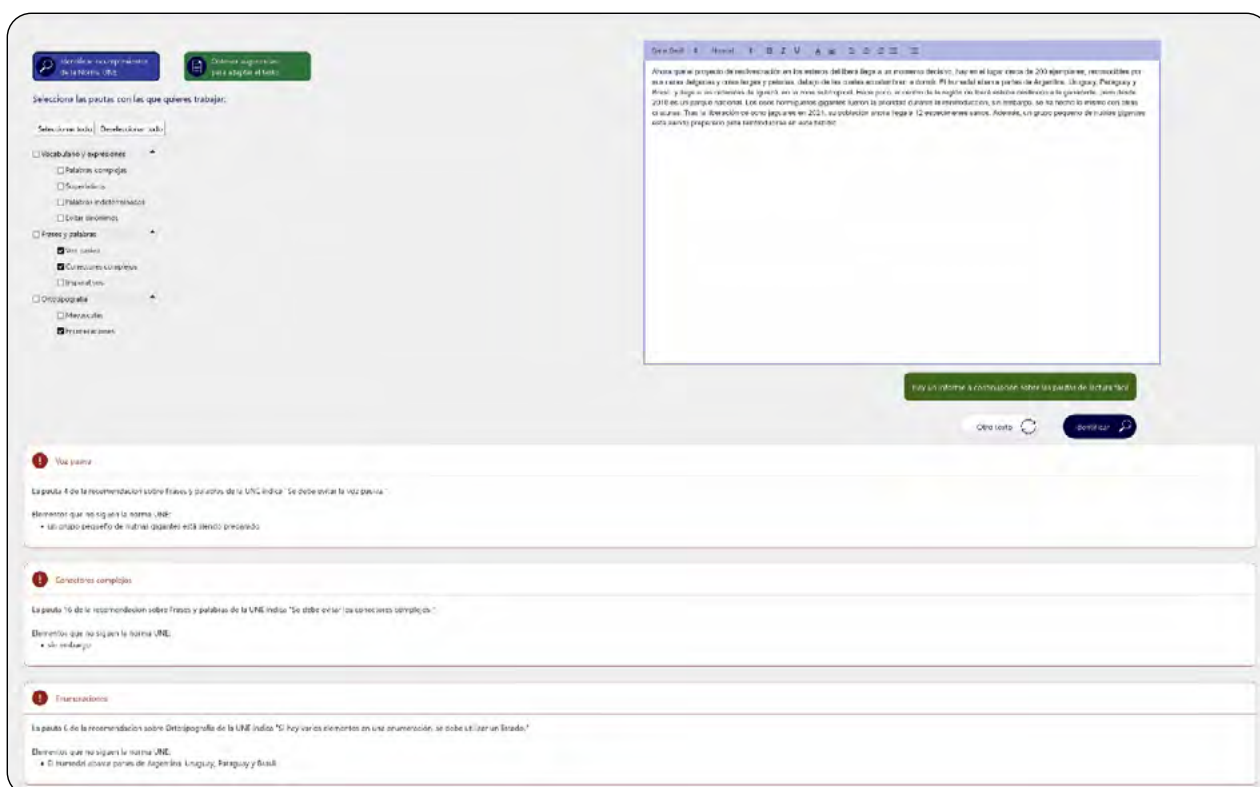


Figura 13. Identificación de voz pasiva, conectores complejos y enumeraciones

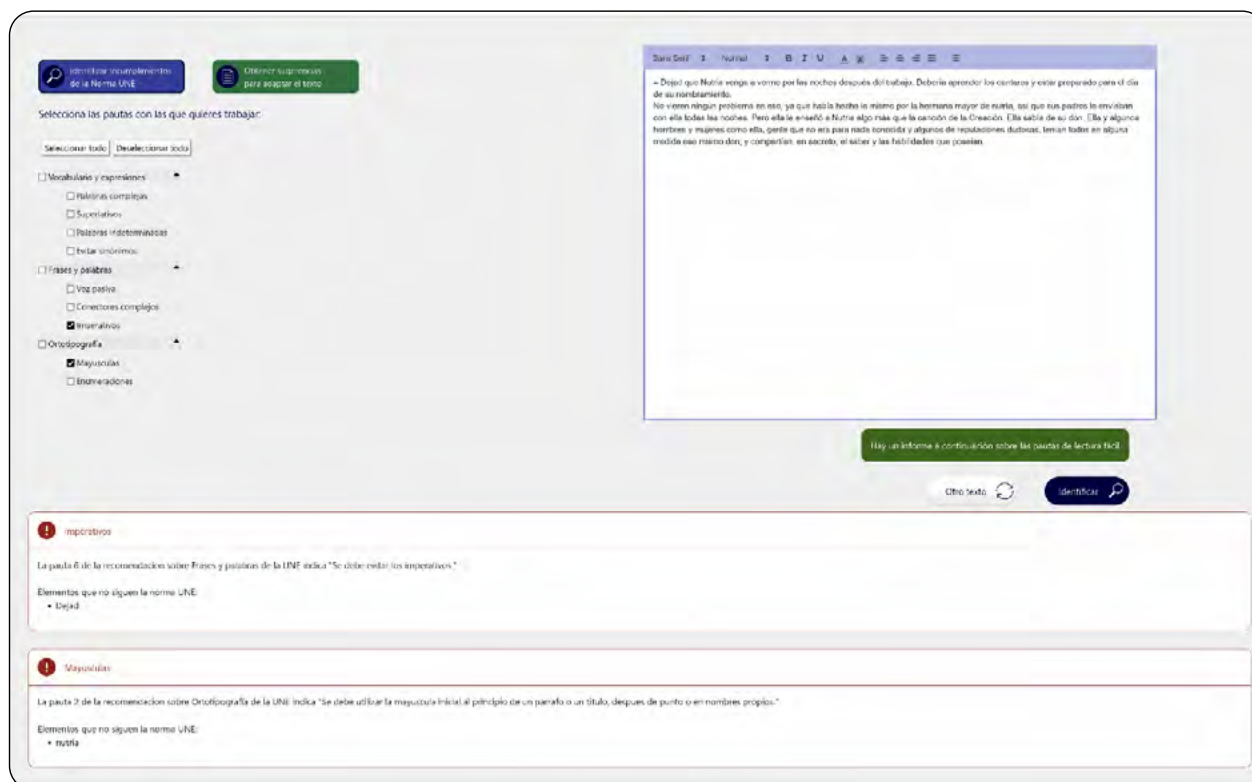


Figura 14. Identificación de imperativos y mayúsculas

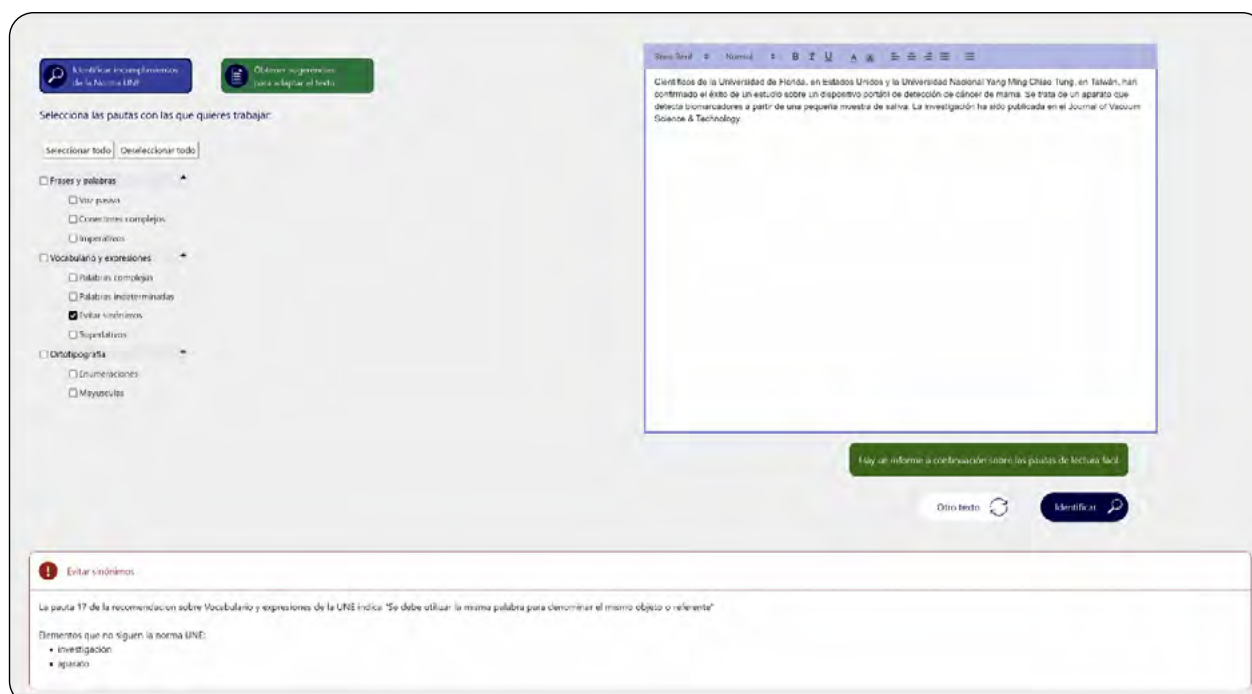
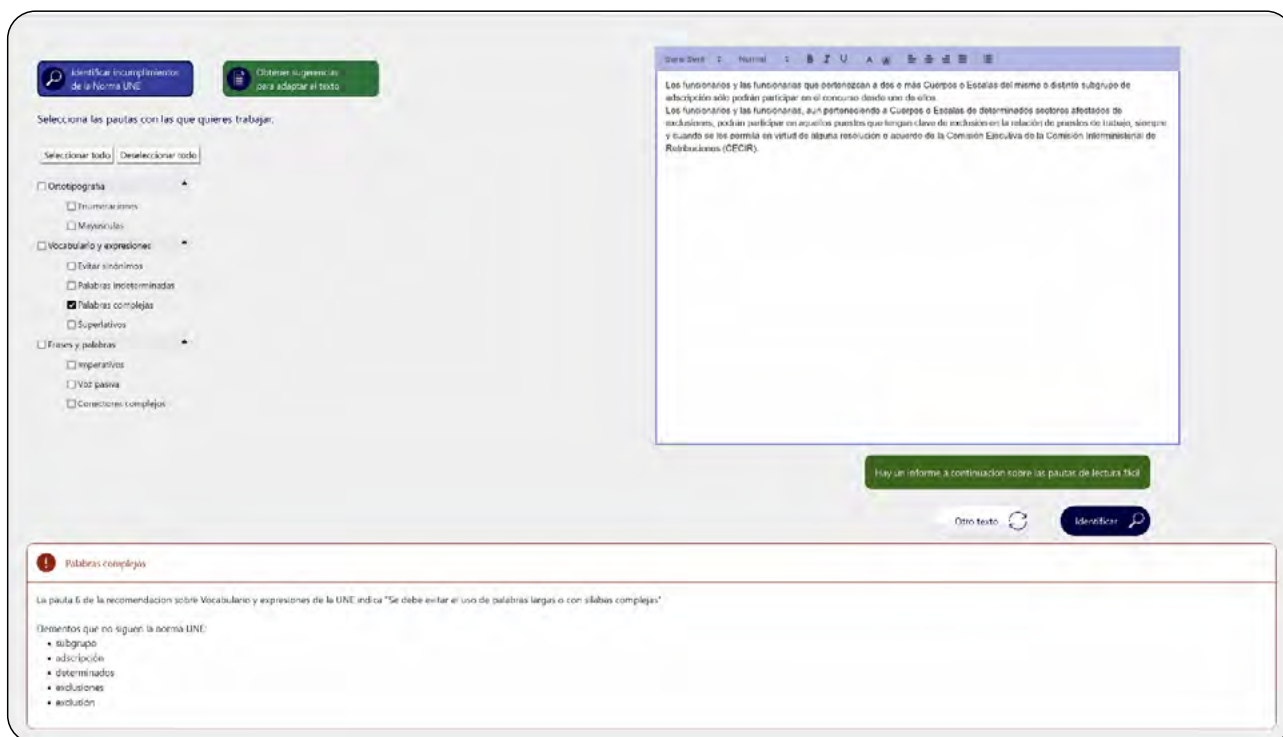


Figura 15. Identificación de sinónimos (ejemplo realizado en una versión futura de la herramienta, no desplegada a fecha del informe)



*Figura 16. Identificación de palabras complejas (ejemplo realizado en una versión futura de la herramienta, no desplegada a fecha del informe)*



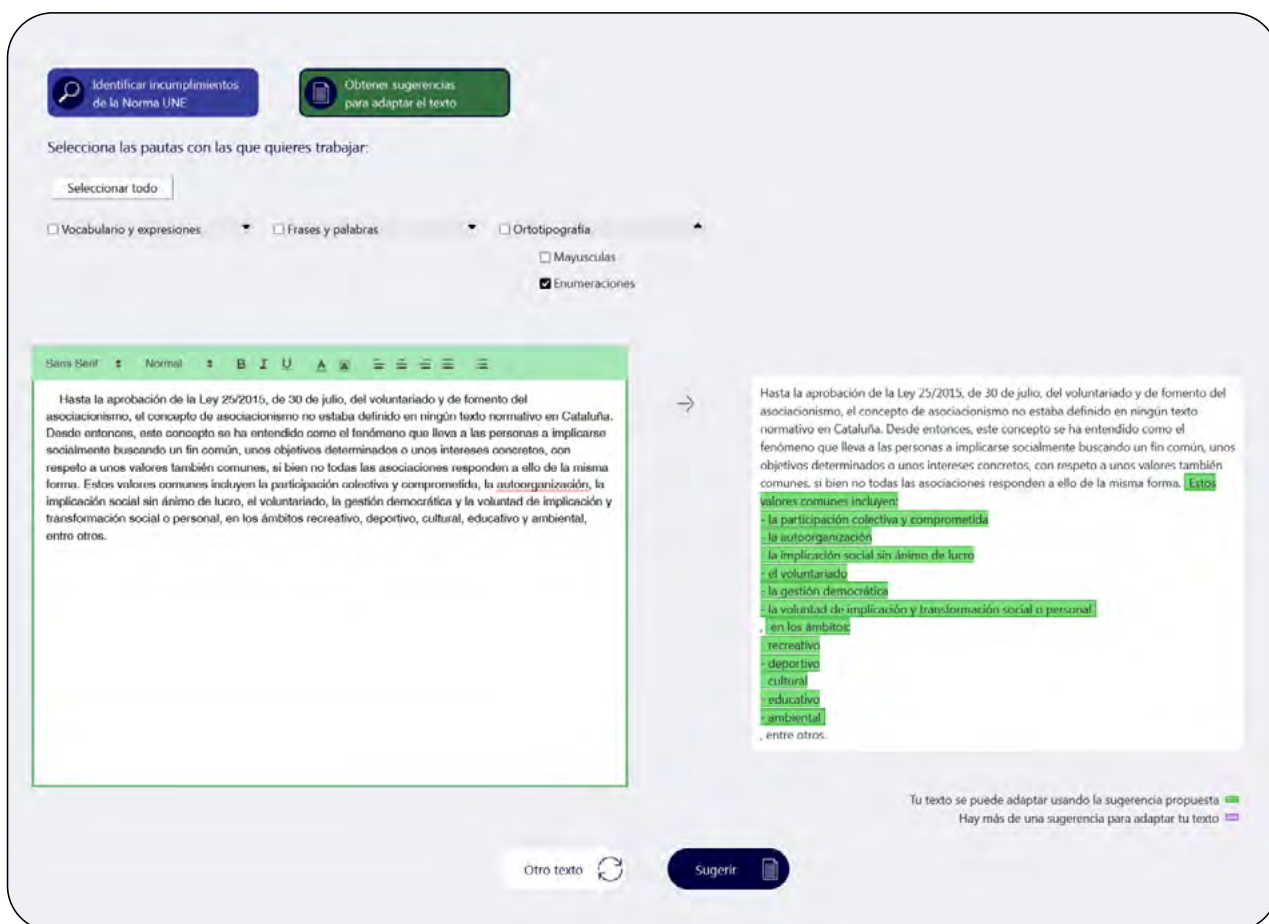


Figura 17. Adaptación de enumeraciones

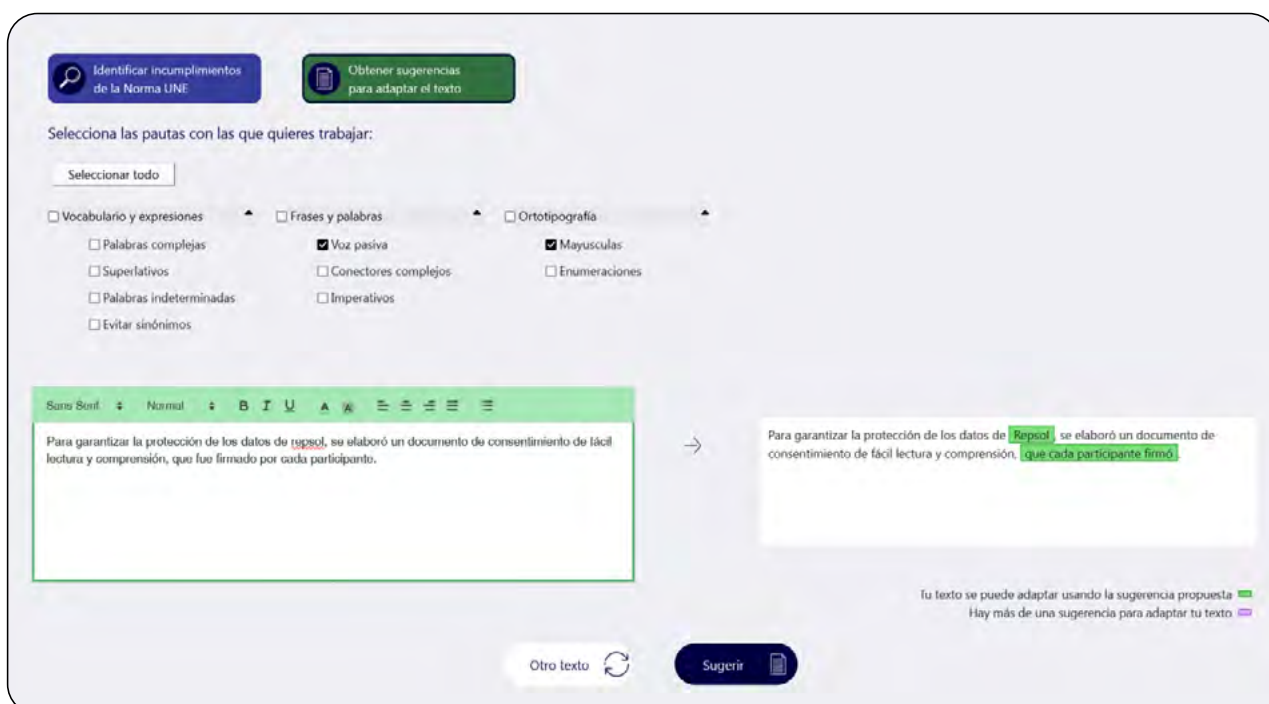


Figura 18. Adaptación de voz pasiva y mayúsculas

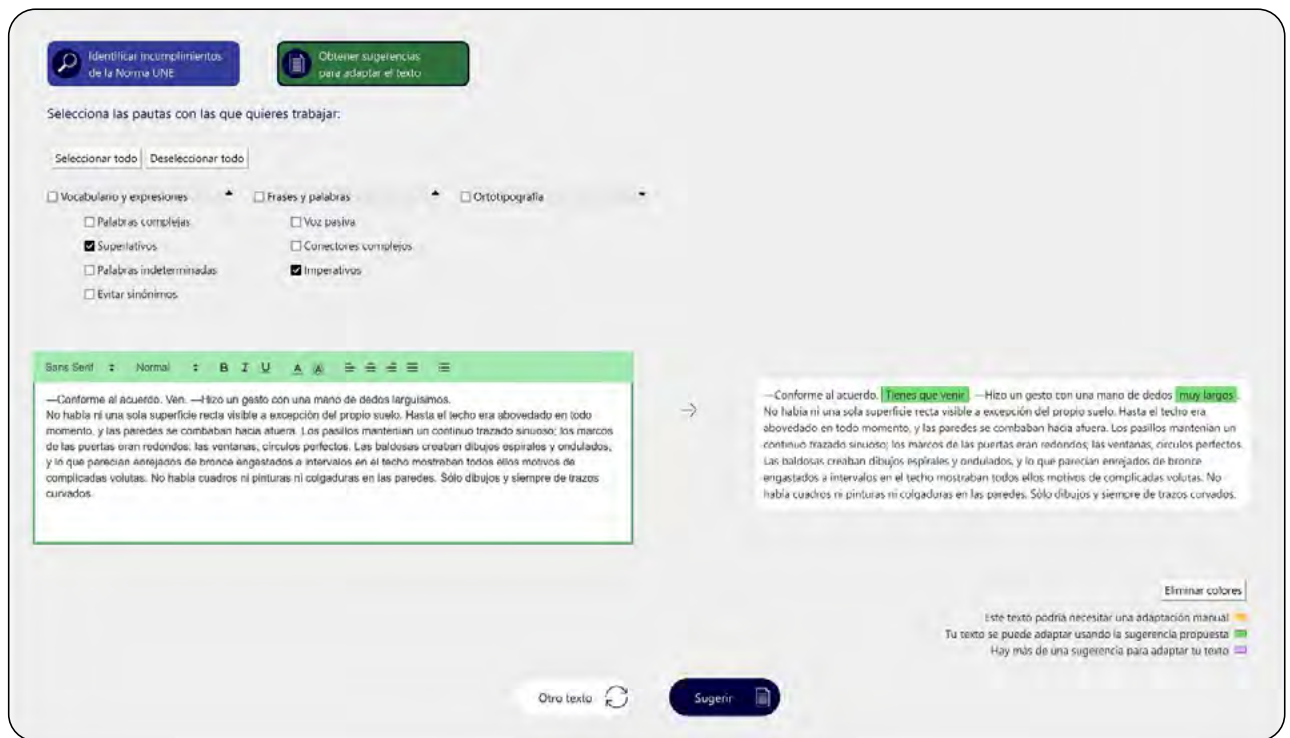


Figura 19. Adaptación de superlativos e imperativos

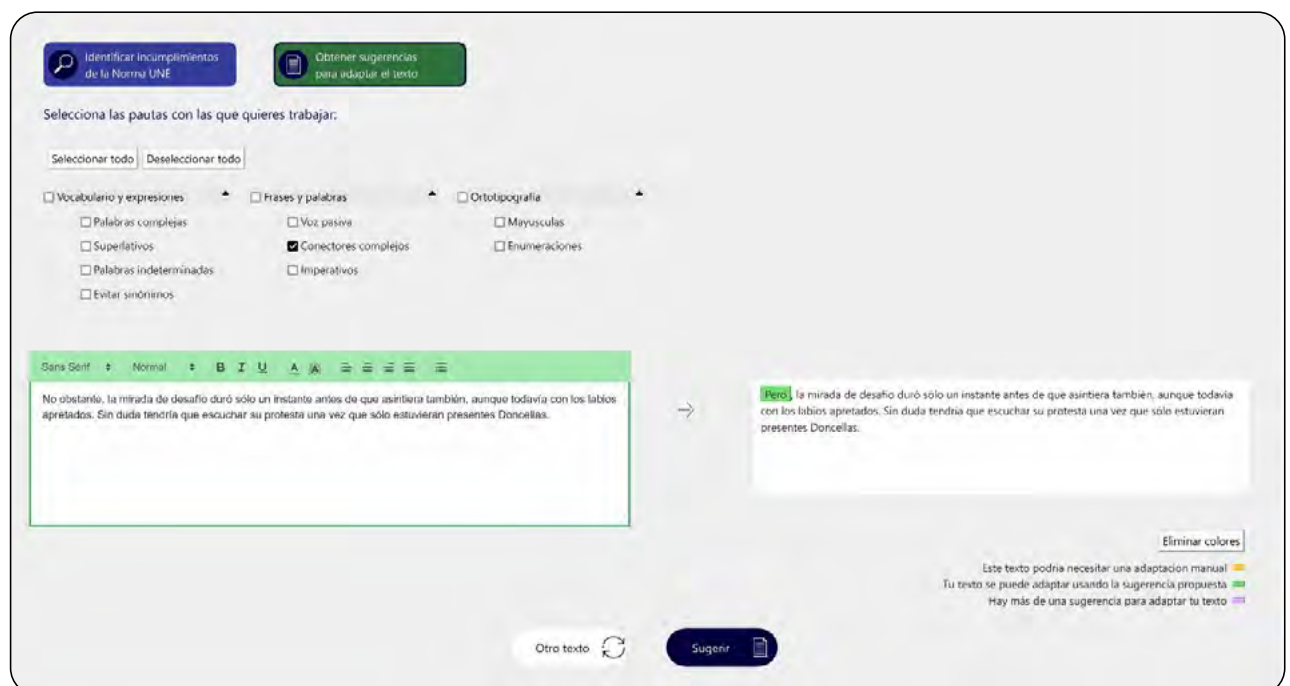


Figura 20. Adaptación de conectores complejos

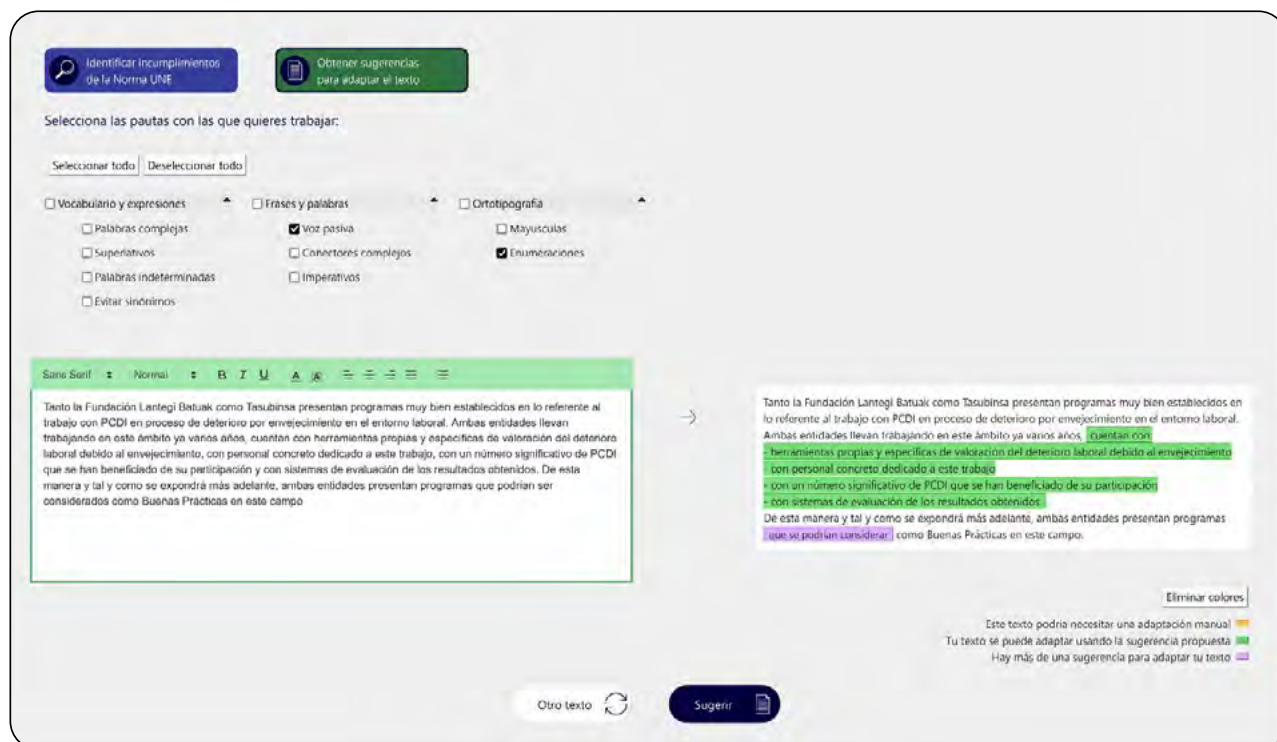


Figura 21. Adaptación de voz pasiva y enumeraciones

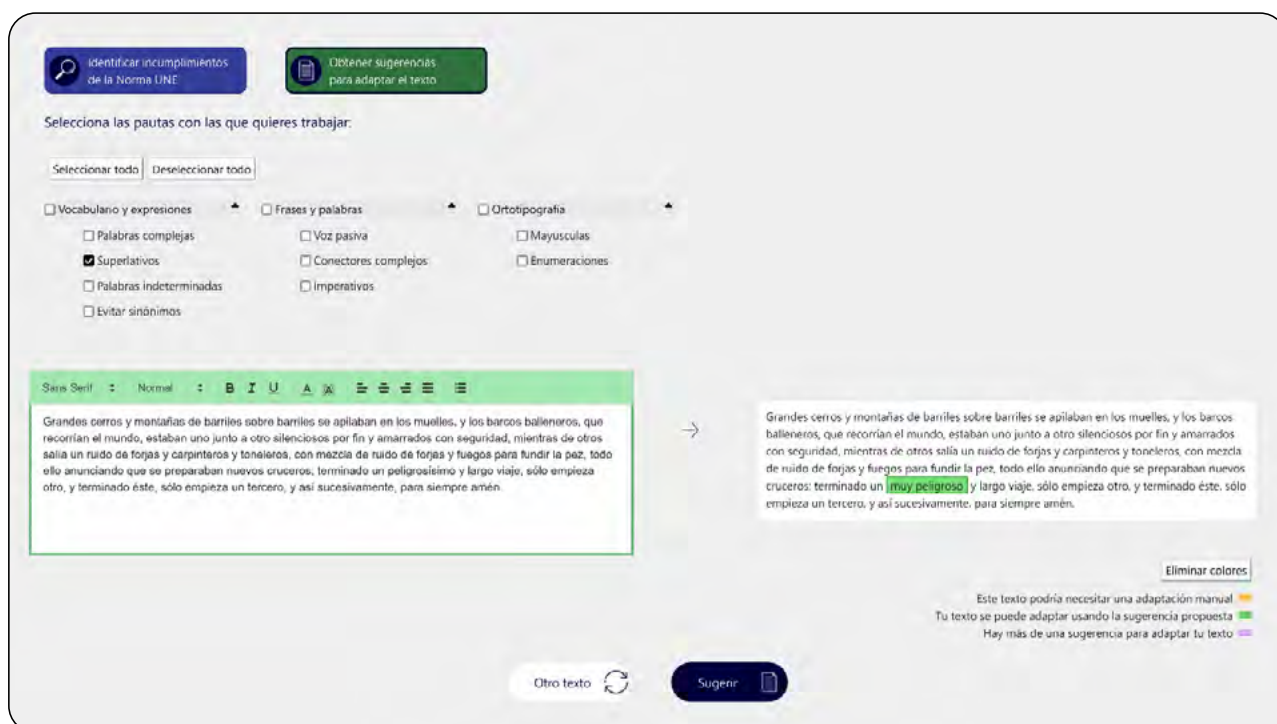


Figura 22. Adaptación de superlativos

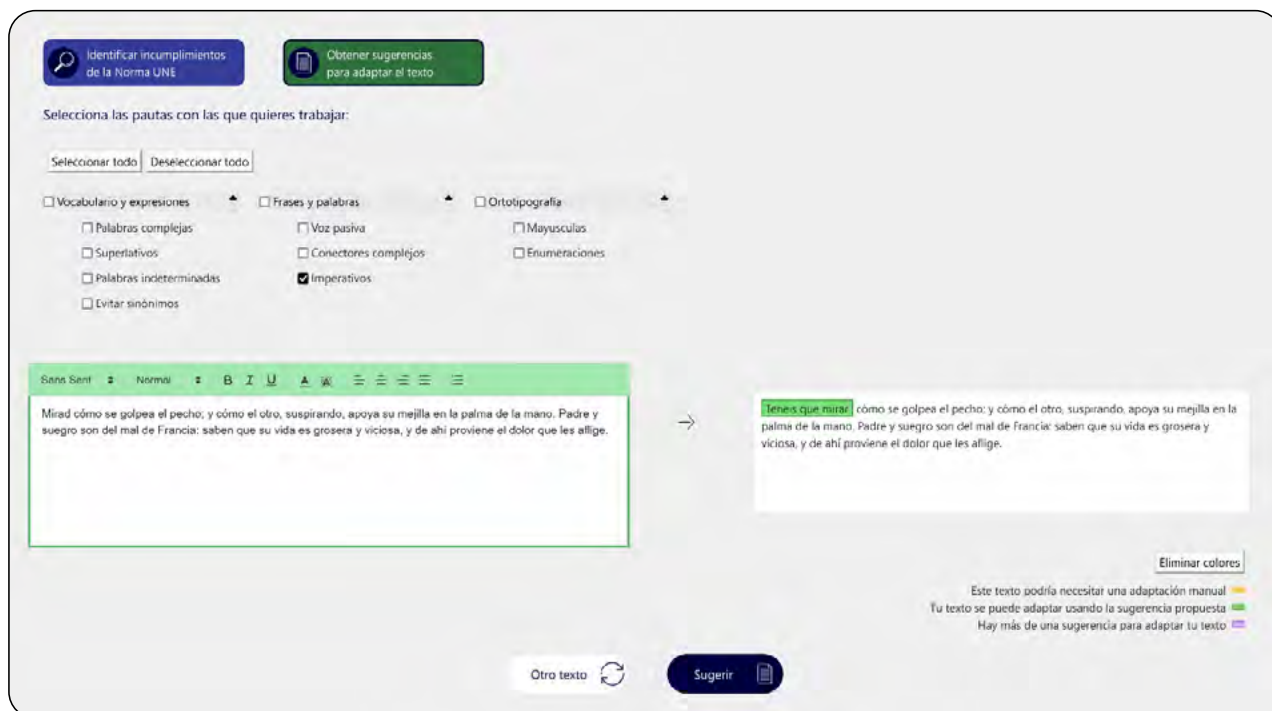


Figura 23. Adaptación de imperativos.

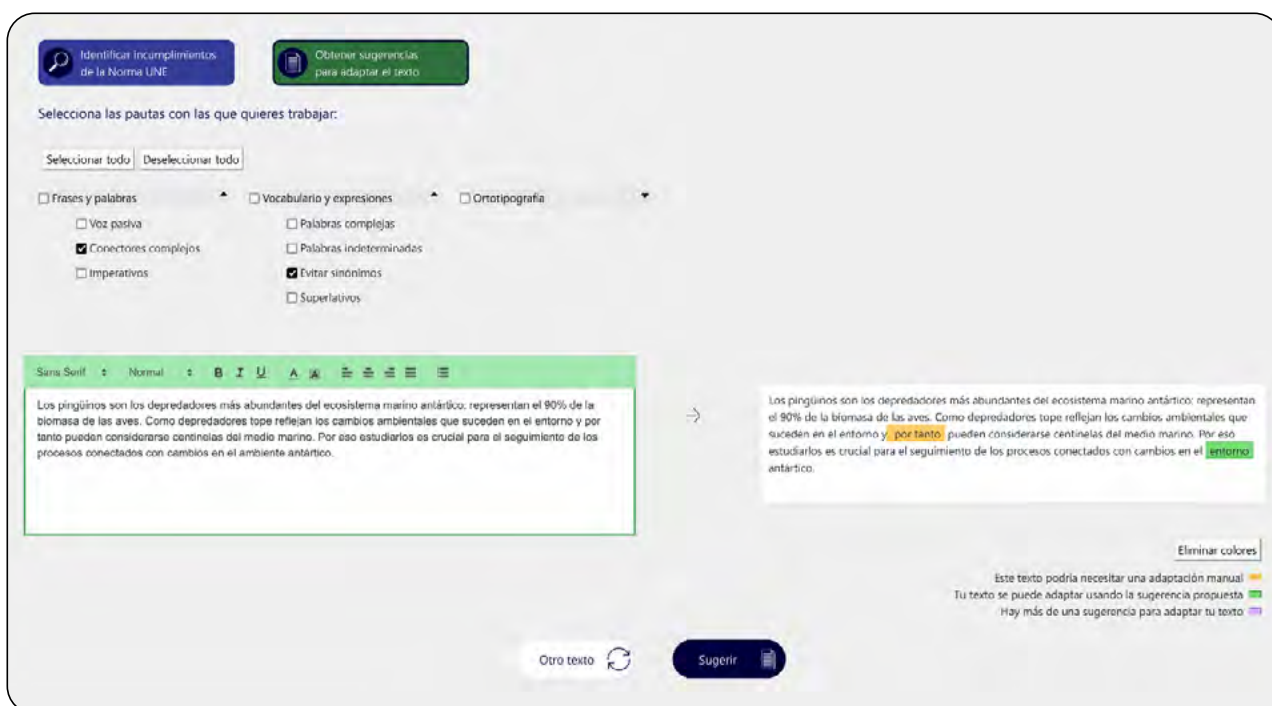


Figura 24. Adaptación de conectores complejos y sinónimos (ejemplo realizado en una versión futura de la herramienta, no desplegada a fecha del informe)

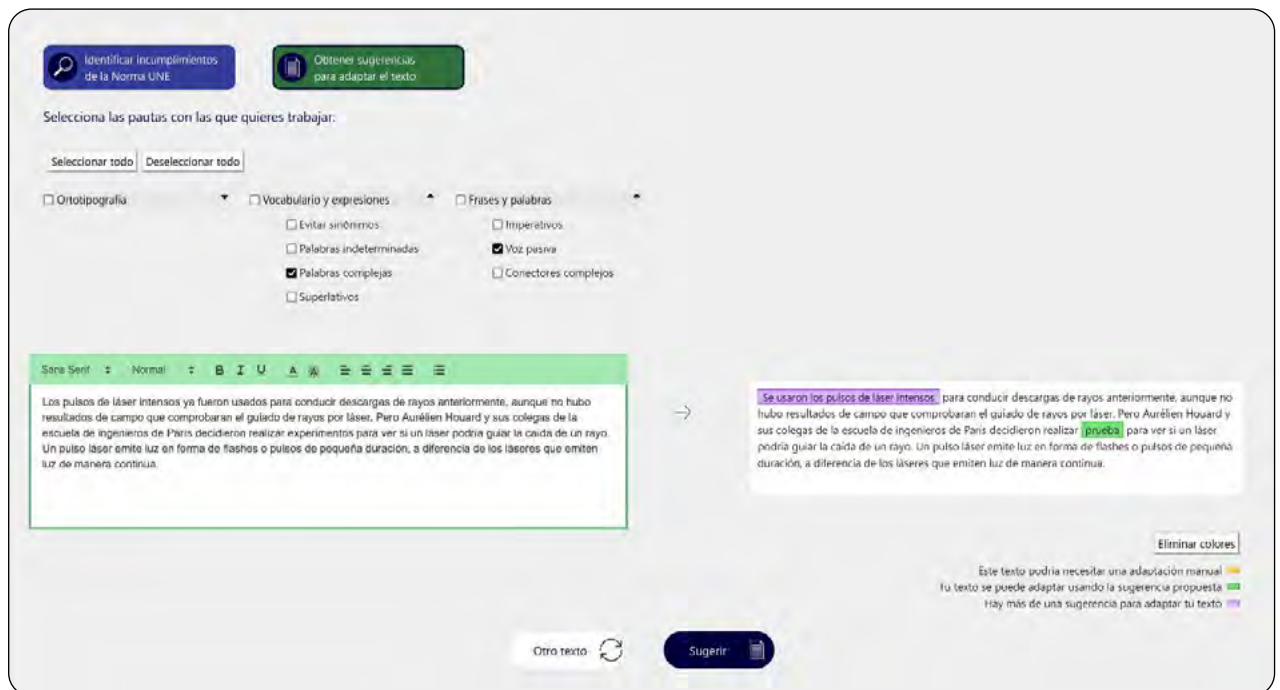


Figura 25. Adaptación de palabras complejas y voz pasiva (ejemplo realizado en una versión futura de la herramienta, no desplegada a fecha del informe)

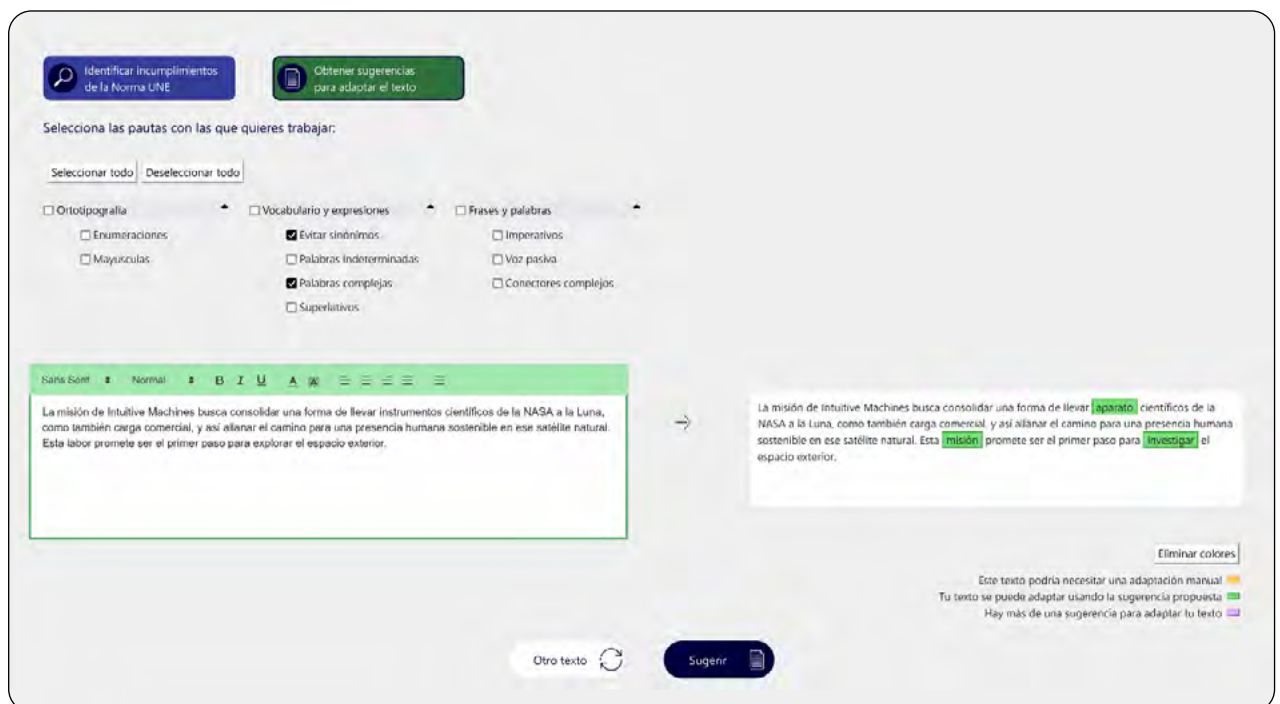


Figura 26. Adaptación de sinónimos y palabras complejas (ejemplo realizado en una versión futura de la herramienta, no desplegada a fecha del informe)



UNIVERSIDAD  
POLITÉCNICA  
DE MADRID

