FACULTAD DE TRADUCCIÓN Y DE INTERPRETACIÓN

GRADO EN TRADUCCIÓN E INTERPRETACIÓN

TRABAJO DE FIN DE GRADO
CURSO 2025-2026

Interpretación judicial: de Núremberg a la inteligencia artificial

Maria Escartín Rodríguez
1668370

TUTORA
MARIA GUIOMAR STAMPA GARCIA-ORMAECHEA

Barcelona, 4 de octubre de 2025

UAB
Universitat Autònoma de Barcelona

AGRADECIMIENTOS

Quisiera expresar mi más sincero agradecimiento a mi tutora, María Guiomar Stampa, por su orientación y acompañamiento a lo largo del desarrollo de este Trabajo de Fin de Grado. Su cercanía, disponibilidad y, especialmente, la confianza depositada en mi trabajo han sido fundamentales para sacar adelante este proyecto. Agradezco enormemente su implicación, sus palabras de ánimo y la seguridad que me ha transmitido durante este proceso.

Asimismo, me gustaría agradecer a la profesora Diana Soliverdi por todos los conocimientos transmitidos en el ámbito de la interpretación, disciplina que he tenido la oportunidad de conocer y aprender en gran medida gracias a ella. Su pasión, dedicación y compromiso a la hora de enseñar han despertado en mí una gran admiración por el mundo de la interpretación.

También quisiera agradecer a mi pareja por su comprensión, afecto y apoyo constante, especialmente en los momentos más difíciles, así como por haber estado siempre a mi lado.

Por último, deseo agradecer a mis padres por su apoyo incondicional, su amor y por acompañarme siempre a lo largo de este camino.

DATOS DEL TFG

Título
Interpretación judicial: de Núremberg a la inteligencia artificial
Títol
Interpretació judicial: de Nuremberg a la intel·ligència artificial
Title
Judicial Interpreting: From Nuremberg to Artificial Intelligence

Autora: Maria Escartín Rodríguez
Tutora: Maria Guiomar Stampa Garcia-Ormaechea
Centro: Universidad Autónoma de Barcelona
Estudios: Grado en Traducción e Interpretación
Curso académico: 2025-2026

Palabras clave
Interpretación judicial; Juicios de Núremberg; Interpretación simultánea; Inteligencia artificial; Ética profesional; Tecnología lingüística
Paraules clau
Interpretació judicial; Judicis de Nuremberg; Interpretació simultània; Intel·ligència artificial; Ètica professional; Tecnologia lingüística
Key words
Judicial Interpreting; Nuremberg Trials; Simultaneous Interpreting; Artificial Intelligence; Professional Ethics; Language Technology

Resumen
Este Trabajo de Fin de Grado analiza la evolución de la interpretación judicial, desde la consolidación de la interpretación simultánea durante los Juicios de Núremberg hasta la incorporación de herramientas de inteligencia artificial en el ámbito jurídico contemporáneo. El estudio examina cómo Núremberg supuso un punto de inflexión en la profesionalización de la interpretación judicial, sentando las bases de principios éticos, técnicos y procesales que continúan vigentes en la actualidad. Con el fin de evaluar hasta qué punto la inteligencia artificial puede asumir funciones tradicionalmente desempeñadas por intérpretes humanos, se realiza un estudio de caso basado en el testimonio de Kyle Rittenhouse (2020), mediante el análisis del rendimiento de DeepL Voice en situaciones de intensidad emocional y complejidad sintáctica. Los resultados sugieren que, aunque la inteligencia artificial representa un avance significativo en accesibilidad e inmediatez, su uso autónomo sigue siendo insuficiente en procedimientos penales debido a sus limitaciones para preservar la precisión lingüística y los matices pragmáticos, emocionales y éticos requeridos en la interpretación judicial. El estudio concluye que, por el momento, la inteligencia artificial debe emplearse exclusivamente como herramienta de apoyo profesional, ya que el intérprete humano sigue siendo indispensable para garantizar un juicio justo y el respeto de las garantías procesales.

Resum

Aquest Treball de Fi de Grau analitza l'evolució de la interpretació judicial, des de la consolidació de la interpretació simultània durant els Judicis de Nuremberg fins a la incorporació d'eines d'intel·ligència artificial en l'àmbit jurídic contemporani. L'estudi examina com Nuremberg va suposar un punt d'inflexió en la professionalització de la interpretació judicial, establint les bases de principis ètics, tècnics i processals que continuen vigents en l'actualitat. Amb la finalitat d'avaluar fins a quin punt la intel·ligència artificial pot assumir funcions tradicionalment exercides per intèrprets humans, es duu a terme un estudi de cas basat en el testimoni de Kyle Rittenhouse (2020), mitjançant l'anàlisi del rendiment de DeepL Voice en situacions d'intensitat emocional i complexitat sintàctica. Els resultats suggereixen que, tot i que la intel·ligència artificial representa un avenç significatiu en accessibilitat i immediatesa, el seu ús autònom continua sent insuficient en procediments penals a causa de les seves limitacions per preservar la precisió lingüística i els matisos pragmàtics, emocionals i ètics requerits en la interpretació judicial. L'estudi conclou que, de moment, la intel·ligència artificial s'ha d'emprar exclusivament com a eina de suport professional, ja que l'intèrpret humà continua sent indispensable per garantir un judici just i el respecte de les garanties processals.

Abstract

This Final Degree Project examines the evolution of judicial interpreting, from the consolidation of simultaneous interpreting during the Nuremberg Trials to the incorporation of artificial intelligence tools into contemporary legal settings. The study explores how Nuremberg represented a turning point in the professionalization of judicial interpreting, laying the foundations for ethical, technical, and procedural principles that remain relevant today. In order to assess the extent to which artificial intelligence may assume functions traditionally performed by human interpreters, a case study based on the testimony of Kyle Rittenhouse (2020) is conducted through the analysis of DeepL Voice performance under conditions of emotional intensity and syntactic complexity. The findings suggest that, although artificial intelligence represents a significant advance in terms of accessibility and immediacy, its autonomous use remains insufficient in criminal proceedings due to its limitations in preserving linguistic accuracy and the pragmatic, emotional, and ethical nuances required in judicial interpreting. The study concludes that, for the time being, artificial intelligence should be used exclusively as a professional support tool, as the human interpreter remains indispensable to ensuring a fair trial and respect for procedural guarantees.

TABLA DE CONTENIDO

1. INTRODUCCIÓN 1

1.1. PRESENTACIÓN DEL TEMA 1
1.2. JUSTIFICACIÓN Y MOTIVACIONES 2
1.3. OBJETIVOS Y ESTRUCTURA 2

2. LOS JUICIOS DE NÚREMBERG Y EL NACIMIENTO DE LA INTERPRETACIÓN SIMULTÁNEA 3

2.1. CONTEXTO HISTÓRICO Y NECESIDAD MULTILINGÜE 3
2.2. LIMITACIONES DE LA INTERPRETACIÓN CONSECUTIVA 4
2.3. LA TECNOLOGÍA IBM Y LA CREACIÓN DE LA SIMULTÁNEA MODERNA 6
2.4. DESAFÍOS COGNITIVOS Y DILEMAS ÉTICOS: EL FACTOR HUMANO 7
2.4.1. APTITUD COGNITIVA Y SELECCIÓN: LA PARADOJA DEL EXPERTO 7
2.4.2. EL INTÉRPRETE COMO "PIEDRA": EL CONFLICTO ENTRE ÉTICA Y FIDELIDAD 7
2.4.3. COMPLICIDAD Y DIPLOMACIA: EL PODER EN LA SOMBRA 8
2.4.4. EL IMPACTO EN EL RITMO DE LA JUSTICIA 8

3. LA INTERPRETACIÓN EN LA ACTUALIDAD 9

3.1. MODALIDADES ACTUALES: SIMULTÁNEA, CONSECUTIVA Y JUDICIAL 9
3.2. COMPETENCIAS DEL INTÉRPRETE PROFESIONAL 10
3.3. REQUISITOS EN CONTEXTOS JURÍDICOS 11

4. INTELIGENCIA ARTIFICIAL APLICADA A LA INTERPRETACIÓN 13

4.1. TECNOLOGÍAS ACTUALES: RECONOCIMIENTO DE VOZ, TRADUCCIÓN AUTOMÁTICA Y LLM 13
4.1.1. EVOLUCIÓN DEL RECONOCIMIENTO DE VOZ (ASR) 13
4.1.2. DE LOS SISTEMAS EN CASCADA A LOS MODELOS "END-TO-END" 13
4.1.3. INTERPRETACIÓN ASISTIDA POR ORDENADOR (CAI) Y APOYO EN CABINA 14
4.1.4. IA GENERATIVA Y PLATAFORMAS DE INTERPRETACIÓN AUTÓNOMA 14
4.2. IMPACTO DE LA IA EN LA INTERPRETACIÓN: BENEFICIOS Y DESAFÍOS CRÍTICOS 14
4.3. LIMITACIONES EN EL CONTEXTO LEGAL: LA ALTERACIÓN DE LA EVIDENCIA 15
4.4. MARCO ÉTICO Y REQUISITOS LEGALES: LA RESPONSABILIDAD Y LA "CAJA NEGRA" 15

5. EL DEBATE ENTRE LA AUTOMATIZACIÓN DIGITAL Y EL JUICIO HUMANO EN EL ÁMBITO PENAL 16

5.1. EL "EFECTO CASCADA" Y LOS RIESGOS COGNITIVOS EN EL TRIBUNAL 16
5.2. EL CRITERIO INSTITUCIONAL DE LA AIIC Y LOS ESCENARIOS DE ALTO RIESGO JUDICIAL 18
5.3. LA IA COMO ASISTENTE DE CABINA Y EL CONTROL DEL INTÉRPRETE 19

6. ANÁLISIS PRÁCTICO DE LA INTERPRETACIÓN AUTOMÁTICA EN SEDE JUDICIAL 20

6.1. METODOLOGÍA DE LA INVESTIGACIÓN 21
6.2. ANÁLISIS DE RESULTADOS: DE LA TÉCNICA A LA ÉTICA PROCESAL 22
6.2.1. ANÁLISIS LINGÜÍSTICO 24
6.2.2. RENDIMIENTO TÉCNICO *DEEPL VOICE* 26
6.2.3. DIMENSIÓN EMOCIONAL Y ÉTICA PROCESAL 26

7. CONCLUSIONES 28

7.1. LIMITACIONES DEL ESTUDIO Y PERSPECTIVAS DE FUTURO 30

REFERENCIAS 32

ANEXOS 34

ANEXO I: TRANSCRIPCIONES 34
ANEXO II: CÓDIGO DE ACCESO Y REFERENCIA MULTIMEDIA 37

ÍNDICE DE TABLAS

Tabla 1: ANÁLISIS COMPARATIVO DEL DISCURSO ORIGINAL, SALIDA DE DEEPL VOICE Y REFERENCIA HUMANA. _____ *24*

1. INTRODUCCIÓN

1.1. Presentación del tema

En un contexto histórico marcado por la Segunda Guerra Mundial tienen lugar los juicios de Núremberg. Juicios organizados por los Aliados con el objetivo de juzgar los crímenes del régimen nazi que, junto a sus participantes provenientes de diversos países y lenguas del mundo, constituyeron un acontecimiento fundamental en la historia de la interpretación: la consolidación de la interpretación simultánea. Si bien es cierto que ya se habían realizado intentos previos de interpretación simultánea, no lograron consolidarse debido a limitaciones técnicas y organizativas (Pöchhacker, 2016, p. 29). La magnitud de los trece juicios evidenció la necesidad de contar con un sistema de comunicación preciso y simultáneo en los procedimientos jurídicos que contasen con múltiples lenguas y nacionalidades. Fue en Núremberg donde se sentaron las bases de la interpretación simultánea moderna (Baigorri-Jalón, 2000, p. 270), pero más allá de este aspecto técnico, se encuentra el aspecto ético: intérpretes que debían traducir testimonios de crímenes atroces; sentencias de muerte; condiciones laborales inimaginables actualmente en el oficio, más de 10h sin descanso, entre otros (Gaiba, 1998, p. 101).

A partir de ese momento, con el tiempo, la interpretación simultánea fue evolucionando y profesionalizándose. En este proceso desempeñó un papel relevante la Asociación Internacional de Intérpretes de Conferencias (AIIC), fundada en 1953, que contribuyó a la estandarización de la profesión mediante la elaboración de recomendaciones y normas de trabajo. (Pöchhacker, 2016, p. 30). De manera que se empezaron a establecer normas y protocolos de trabajo: se aislaron cabinas para mejorar la concentración, se implementaron límites de tiempo de trabajo para garantizar la salud de los intérpretes y, en muchos casos, se incorporó un sistema de doble intérprete para turnarse, mejorando así la precisión, garantizando descansos, y ayudando a solventar imprevistos (Baigorri-Jalón, 2000, p. 189). Asimismo, AIIC promovió un código ético profesional basado en principios como la confidencialidad, la fidelidad al discurso original y la imparcialidad del intérprete. Estas mejoras exponen la consolidación de la interpretación como una profesión organizada y que, en el caso de la interpretación judicial, no solo exige habilidades lingüísticas, sino también un compromiso ético con la justicia.

Hoy en día, el auge de la inteligencia artificial plantea nuevas preguntas: ¿sería ético utilizar sistemas de IA en casos judiciales sensibles? ¿Hasta qué punto la inteligencia artificial podría asumir funciones del intérprete humano manteniendo la exactitud lingüística y respetando los

estándares éticos? El presente TFG explorará estas cuestiones, evaluando hasta qué punto la inteligencia artificial puede complementar o sustituir al intérprete sin comprometer la calidad de la interpretación judicial. Estos antecedentes históricos serán la base para realizar en la parte práctica la comparación entre la inteligencia artificial y la interpretación humana.

1.2. Justificación y motivaciones

A lo largo de mi vida, como cualquier otra persona, me han hecho las clásicas preguntas sobre el futuro: "¿Qué quieres ser de mayor?" en la infancia, "¿Qué quieres estudiar?" durante la adolescencia, e incluso "¿A qué quieres dedicarte?" durante la carrera universitaria. Mis respuestas han ido cambiando con el tiempo: quise ser arqueóloga, historiadora, actriz, maestra, abogada, y también consideré estudiar Filología Francesa.

Sin embargo, siempre he sentido una especial predilección por los idiomas. Desde pequeña, se me dieron bien, me gustaban y destacaba en ellos. Pero el descubrimiento de la traducción representó para mí un paso más allá, ya que me fascina la posibilidad de hacer que varias lenguas, varias culturas y, en cierto modo, varios países se comprendan entre sí.
A la par, siempre me ha interesado la justicia, la ley y los juicios, especialmente en el ámbito penal. Pero no quiero limitar un trabajo de fin de grado a la historia o a la ética de los juicios, no lo considero suficiente. Quiero profundizar, investigar, aprender el porqué de los procesos y las decisiones que los rodean.

Actualmente, la inteligencia artificial nos acompaña en gran parte de los ámbitos de la vida, incluida la traducción. Recuerdo que hace algún tiempo, una docente de una universidad española, que impartía clases de lengua extranjera a estudiantes de Traducción e Interpretación, me dijo: "¿Para qué estudias esta lengua? Para interpretar no lo creo: la inteligencia artificial ya es capaz de ello y no te va a servir de nada." Aunque no comparto en absoluto su opinión, su comentario me hizo reflexionar y me impulsó a investigar.
Poco después, asistí a una exposición sobre los Juicios de Núremberg en el Goethe-Institut de Madrid, en la que realicé una visita guiada privada organizada por los profesores de interpretación para mí y mis compañeros. Quedé completamente fascinada y supe en ese momento que debía remontarme a los inicios de la interpretación para desarrollar mi TFG y mi investigación.

1.3. Objetivos y estructura

El presente TFG pretende explorar el origen de la interpretación, narrar su evolución y situarse en la actualidad, con la inteligencia artificial como protagonista. Se analizarán cuestiones sustanciales: ¿realmente sería fiable la interpretación automática en un juicio penal, en un contexto donde no es admisible cometer errores? Éticamente, los intérpretes se ahorrarían la

carga emocional de traducir sentencias y testimonios devastadores, pero ¿puede una máquina, sin emociones y sensibilidad lingüística, actuar con el mismo tacto y comprensión que un ser humano?

Para responder a estas cuestiones, este TFG incluirá un marco teórico, donde se expondrá la información necesaria para comprender el tema en su conjunto; y una parte práctica en la que se utilizarán fragmentos de un juicio para comparar la interpretación humana con la realizada por inteligencia artificial. Solo a través de la práctica será posible evaluar las diferencias, ventajas y limitaciones de cada una de ellas.

2. LOS JUICIOS DE NÚREMBERG Y EL NACIMIENTO DE LA INTERPRETACIÓN SIMULTÁNEA

2.1. Contexto histórico y necesidad multilingüe

Tras el final de la Segunda Guerra Mundial, Europa se encontraba devastada y resultaba esencial juzgar a los responsables del régimen nazi. Para ello, los Aliados —Estados Unidos, Reino Unido, Francia y la Unión Soviética— organizaron los juicios de Núremberg, destinados a juzgar a los principales líderes nazis, entre ellos Hermann Göring, uno de los ayudantes más destacados de Hitler. Las sentencias incluyeron condenas de muerte, cadenas perpetuas y largas penas de prisión, mientras que algunos acusados optaron por el suicidio al conocer su veredicto o durante el transcurso de los juicios (Gaiba, 1998, p. 29).

Estados Unidos pretendía que estos procesos fueran muy mediáticos, mostrando al mundo la magnitud de los crímenes nazis, con la voluntad de educar sobre la barbarie producida, establecer un precedente de justicia internacional, y deslegitimar a la élite nazi, reconstruyendo así los principios éticos tras la guerra. Esta dimensión internacional y pública del juicio hizo imprescindible la coordinación lingüística y la interpretación simultánea, para garantizar que los distintos idiomas y audiencias pudieran comprender los procedimientos en tiempo real (Baigorri-Jalón, 2000, p. 272).

La trascendencia internacional del juicio suponía una dificultad lingüística considerable, dado que los participantes hablaban distintos idiomas y provenían de varios países. Las lenguas en las que se interpretó fueron el inglés, el ruso, el alemán y el francés. En vista de que el juicio se preveía largo, se requirió un gran número de intérpretes. Sin embargo, esta necesidad no solo respondía a la diversidad lingüística, sino también a la duración de los juicios, que dificultaban enormemente la interpretación consecutiva, y al interés de Estados Unidos por garantizar que el proceso se siguiera de manera mediática y fluida.

Es por todo esto por lo que se recurrió a la interpretación simultánea, que ya había sido ensayada anteriormente en entornos comerciales y experimentales, aunque hasta entonces no se había podido consolidar por limitaciones técnicas y organizativas.

La selección de intérpretes se realizó entre ciudadanos con conocimientos de varias lenguas, aunque pocos contaban con formación profesional en traducción o interpretación, escasa en la época, de manera que la mayoría aprendió sobre la marcha. La coordinación del equipo estuvo a cargo del coronel León Dostert, quien supervisó la instalación de las cabinas utilizadas y el uso de sistemas desarrollados en colaboración con IBM (1998, p. 30), acuerdo que se cerró teniendo en cuenta que, posteriormente, estos sistemas podrían ser utilizados en Estados Unidos para eventos judiciales y diplomáticos. El sistema de IBM conectaba cabinas de intérpretes a una central de control, permitiendo así transmitir simultáneamente los discursos a inglés, francés, ruso y alemán. El sistema resultó exitoso y aseguró la comprensión entre los participantes de distintas nacionalidades.

El trabajo de los intérpretes implicó enormes exigencias éticas y físicas, pues debían traducir testimonios sobre crímenes atroces durante jornadas o sentencias de muerte de más de diez horas, lo que generaba gran estrés e incluso episodios de desmayo. Para ello, había una persona encargada de supervisar el estado de los intérpretes: si necesitaban parar, si sufrían alguna crisis de nervios, entre otros problemas de salud. Se ha mencionado, aunque no existe confirmación histórica, que algunas intérpretes llegaban a utilizar horquillas que utilizaban en el pelo para generar un breve cortocircuito en los equipos y así poder descansar momentáneamente durante las largas jornadas. Sumado a ello, cabe recalcar que algunos intérpretes habían perdido familiares en campos de concentración o habían vivido experiencias muy cercanas (Goethe-Institut, 2025). Estas condiciones reflejan la complejidad del proceso, así como la perspectiva ética y humana de la interpretación judicial, un aspecto que sigue siendo fundamental en la profesión hoy en día.

2.2. Limitaciones de la interpretación consecutiva

Como se ha expuesto anteriormente, antes de la consolidación de la interpretación simultánea, la modalidad que predominaba en contextos formales e institucionales era la interpretación consecutiva. En este sistema, el orador se expresaba durante un período de tiempo y, posteriormente, el intérprete reproducía el discurso en la lengua de llegada, apoyándose en la memoria y en técnicas de toma de notas. Esta modalidad había funcionado eficazmente en congresos, reuniones diplomáticas y negociaciones internacionales, así como en la Sociedad de Naciones, donde el francés y el inglés eran las principales lenguas de trabajo y el número limitado de idiomas y participantes permitía organizar las intervenciones sin grandes complicaciones (Pöchhacker, 2016, p. 29). Sin embargo, los Juicios de Núremberg, con cuatro lenguas oficiales y un elevado número de participantes y testimonios, evidenciaron las claras limitaciones de la consecutiva en un contexto judicial de gran escala, haciendo necesaria la implementación de la interpretación simultánea (Baigorri-Jalón, 2000, p. 270).

Uno de los principales problemas de la interpretación consecutiva era el factor temporal. Cada intervención debía repetirse completamente en las distintas lenguas, lo que alargaba considerablemente la duración de las sesiones. En un juicio con numerosos acusados, testigos, documentos legales, este sistema habría duplicado o incluso triplicado el tiempo necesario para cada audiencia, siendo muy poco viable la interpretación durante el proceso (Baigorri-Jalón, 2000, p. 46**).** Además, la acumulación de retrasos podía interferir con la programación del tribunal.

A ello se sumaba la complejidad del contenido. Los juicios de Núremberg incluían testimonios especialmente extensos sobre crímenes de guerra, informes técnicos y alegatos jurídicos con un alto grado de precisión terminológica. La interpretación consecutiva hubiese exigido que el intérprete retuviese grandes fragmentos de información antes de reproducirlos, lo que aumentaba el riesgo de omisiones, reformulaciones involuntarias o errores de traducción (Pöchhacker, 2016, p. 111). Este riesgo resultaba especialmente relevante en un contexto tan delicado donde cada palabra podía tener consecuencias graves y determinar sentencias.
Otra limitación relevante residía en la carga cognitiva y emocional del intérprete. La consecutiva requiere una concentración prolongada, ya que el intérprete debe escuchar, analizar, memorizar y reproducir el discurso en la lengua de llegada de manera precisa. En los juicios de Núremberg, donde los intérpretes se enfrentaban a relatos de violencia extrema y a largas jornadas sin descanso, este esfuerzo mental hubiese sido insostenible y la carga emocional podía interferir en la reformulación del mensaje.
Asimismo, la interpretación consecutiva presentaba problemas desde el punto de vista de la dinámica judicial. Las interrupciones constantes para permitir la interpretación dificultaban la interacción directa entre jueces, fiscales, abogados y acusados. En un proceso que Estados Unidos deseaba que fuera mediático y accesible para la prensa y la opinión pública, estas pausas reducían el impacto comunicativo (Baigorri-Jalón, 2000, p. 46).

Por último, la consecutiva no respondía a la necesidad multilingüe simultánea de Núremberg. Al existir cuatro lenguas oficiales, cada intervención habría tenido que ser interpretada varias veces sucesivamente, prolongando todavía más el juicio y aumentando la probabilidad de inconsistencias en las versiones. Esto sacaba a la luz que la interpretación consecutiva no estaba diseñada para contextos judiciales de carácter internacional, con gran número de participantes y de gran duración. (Baigorri-Jalón, 2000, p. 272).

Estas limitaciones evidenciaron que, aunque la interpretación consecutiva era válida en otros contextos, resultaba insuficiente en un juicio internacional de estas dimensiones. La necesidad de rapidez, precisión, y acceso simultaneo a la información llevó a replantear el modelo de interpretación utilizado hasta entonces. De este modo, los juicios de Núremberg se convirtieron

en un punto de inflexión que impulsó a que la interpretación simultánea se adoptase como solución más eficaz para contextos judiciales complejos, marcando un antes y un después en la historia de la interpretación.

2.3. La tecnología IBM y la creación de la simultánea moderna

La consolidación de la interpretación simultánea en Núremberg no fue solo un logro organizativo del coronel Dostert, como ya se ha mencionado, sino que fue posible gracias a la adaptación de tecnologías previas. El pilar de este avance fue el sistema *Hushaphone Filene-Finlay*, un equipo diseñado en los años 20 y fabricado por la empresa IBM. Como explica Gaiba (1998, p. 30), este dispositivo funcionaba mediante un complejo sistema de cables y selectores que permitía a cada asistente elegir en sus auriculares entre el discurso original o cualquiera de las cuatro traducciones en tiempo real.

Sin embargo, la puesta en marcha de esta tecnología fue tan difícil como ambiciosa. El material llegó a Núremberg apenas tres semanas antes de que empezara el juicio, en un *Justizpalast* 'Palacio de Justicia' que todavía estaba en ruinas por la guerra. La situación fue tan tensa que, tras el hundimiento de una parte del suelo de la sala principal durante las obras, el equipo técnico tuvo que realizar una instalación provisional en el ático antes de pasar a la definitiva (Gaiba, 1998, p. 39). Además de los problemas del edificio, faltaba material: inicialmente solo había 200 auriculares, lo que obligó a buscar de urgencia otras 300 unidades en Ginebra para cubrir a todo el tribunal (Gaiba, 1998, p. 40). Para que todo funcionara, la propia IBM envió a sus ingenieros para trabajar junto a los técnicos del ejército estadounidense, bajo la dirección del Mayor Vincent, quien ya conocía la empresa por haber trabajado en ella anteriormente (Gaiba, 1998, p. 39).

Más allá de los cables y cascos, la tecnología de IBM cambió por completo la jerarquía en la sala. Según explica Baigorri Jalón (2000, pp. 289-290), el intérprete dejó de ser una figura invisible para convertirse en el regulador del ritmo del juicio mediante un sistema de señales luminosas. Ante la incapacidad de procesar el discurso más allá de cierta velocidad, surgió la idea de instalar un dispositivo con dos luces que el monitor accionaba para avisar al orador: una luz amarilla para pedir que se hablara más despacio y una luz roja para detener por completo la intervención (Baigorri Jalón, 2000, p. 289). Esta "autoridad del interruptor" otorgó a las cabinas un control sobre el tiempo judicial que nunca antes había existido.

Finalmente, este despliegue técnico obligó a Dostert a organizar el trabajo de forma más humana para evitar errores graves causados por el cansancio. Al ser una tarea mentalmente agotadora, se instauró un sistema de rotación por equipos "regulado como papel de música" (Baigorri Jalón, 2000, p. 301). Se establecieron tres turnos (A, B y C) de doce intérpretes cada uno, de modo que el tiempo de interpretación continuada no excediera los 45 minutos. Mientras

un equipo trabajaba en las cabinas, el segundo permanecía en una sala contigua (la sala 606) escuchando los debates para no perder el hilo terminológico, y el tercero descansaba (Baigorri Jalón, 2000, pp. 301-302). Con esta organización se sentaron las bases de las normas de salud laboral y rendimiento que, todavía hoy, rigen la interpretación profesional.

De este modo, la unión entre la infraestructura técnica de IBM y la nueva organización profesional permitió que la interpretación simultánea dejara de ser un experimento para convertirse en el estándar de la justicia internacional.

2.4. Desafíos cognitivos y dilemas éticos: el factor humano

El éxito de la interpretación en Núremberg no dependió únicamente de la tecnología de IBM, sino de la resistencia psicológica y la calidad ética de un grupo de profesionales sometidos a una presión sin precedentes. Este apartado analiza cómo la personalidad y la capacidad cognitiva de los intérpretes moldearon el curso del juicio.

2.4.1. Aptitud cognitiva y selección: la paradoja del experto

Al no existir criterios previos de calidad para la interpretación simultánea, la selección de personal fue un experimento en sí mismo. Los examinadores buscaban algo más que un simple bilingüismo y conocimiento riguroso de las lenguas de trabajo: se requería "compostura y capacidad de mantener la calma en situaciones estresantes" (Gaiba, 1998, p. 46). Según Skinner y Carson (1990, p. 15), este proceso reveló una paradoja: los candidatos con perfiles puramente académicos y amplios conocimientos de esos idiomas a menudo fracasaban en la cabina. Su búsqueda de la perfección les impedía adaptarse a la naturaleza inmediata de la simultánea, donde no se puede retroceder. La interpretación requería una agilidad mental para "escuchar y hablar al mismo tiempo" y tomar decisiones instantáneas bajo presión física (Gaiba, 1998, p. 47). Esta capacidad se apoya en lo que Marianne Lederer (1990, pp. 53-54) denomina "complementos cognitivos": el intérprete no solo transcodifica sonidos, sino que utiliza su conocimiento del mundo y del contexto para dar sentido al discurso. En Núremberg, entender la terminología nazi era tan vital como el idioma; sin estos complementos, la traducción perdía sus implicaciones legales (Lederer, 1990, p. 54).

2.4.2. El intérprete como "piedra": el conflicto entre ética y fidelidad

Uno de los mayores desafíos éticos fue la tensión entre la invisibilidad profesional y la humanidad. W. Brian Altano (1990, p. 96) describe la presión de los tribunales que exigían que el intérprete actuara como una "máquina" o una "piedra". Sin embargo, la voz del intérprete podía alterar el testimonio mediante la inflexión vocal (Gaiba, 1998, p. 106). Existieron casos

de personalidades opuestas: desde Margot Bortlin, famosa por su interpretación dramática y gestual, hasta intérpretes que resumían tanto las frases que confundían a la audiencia.

Este factor humano llegó a comprometer la exactitud legal cuando el pudor o el rechazo moral interferían. Gaiba (1998, p. 107) relata cómo algunas intérpretes se negaron a traducir lenguaje soez o suavizaron insultos dirigidos a las víctimas (como traducir "mearse en los judíos" por "ignorar a los judíos"). En estos casos, las intérpretes fueron reemplazadas por violar el principio de fidelidad, recordándoles que eran "servidores de la corte" obligados a la precisión absoluta frente a lo moralmente repugnante (Gaiba, 1998, p. 108). Esta exigencia no era solo ética, sino que respondía al mandato legal del Estatuto del Tribunal recogido en el Volumen 1 (IMT, 1947). Según el Artículo 25, todos los procedimientos debían conducirse obligatoriamente en las lenguas oficiales y en la del acusado (p. 20). Por tanto, cualquier alteración del sentido por parte del intérprete suponía una violación directa del derecho del acusado a comprender el proceso, poniendo en riesgo la validez jurídica de las actuaciones bajo el pretexto de una falsa cortesía o pudor personal.

2.4.3. Complicidad y diplomacia: el poder en la sombra

El intérprete se encontraba "posado precariamente entre la imparcialidad profesional y un vínculo complejo con el testigo" (Altano, 1990, p. 100). En la cercanía de las cabinas, surgió una extraña "complicidad lingüística": acusados como Albert Speer ayudaban a los intérpretes pasándoles notas con términos técnicos difíciles (Gaiba, 1998, p. 129). Este vínculo humano contrastaba con la responsabilidad geopolítica del cargo. Según Ruth Roland (1999, p. 4), los intérpretes poseían un poder intrínseco que emanaba de su credibilidad, actuando como "diplomáticos lingüísticos".

El miedo a que una mala interpretación provocara roces diplomáticos entre los Aliados era constante. Como ocurrió históricamente con las discrepancias textuales en tratados previos (Roland, 1999, pp. 124-125), un error de matiz en Núremberg podía cambiar el sentido de una prueba judicial. El intérprete debía navegar entre la lealtad a la verdad y la necesidad de mantener la cohesión comunicativa entre jueces con visiones legales contrapuestas.

2.4.4. El impacto en el ritmo de la justicia

Por último, la interpretación alteró la estrategia judicial. Los fiscales se quejaban de que la simultánea eliminaba el ritmo de los interrogatorios cruzados, dando a los acusados tiempo extra para pensar sus respuestas mientras esperaban la traducción (Gaiba, 1998, p. 102). Figuras como Hermann Göring explotaron esta debilidad del sistema para ganar ventaja psicológica, demostrando que la interpretación no fue solo un puente lingüístico, sino un factor que influyó

en la equidad y el desarrollo del propio juicio (Gaiba, 1998, p. 101). Esta ralentización del ritmo procesal entraba en conflicto directo con las facultades del Tribunal establecidas en el Artículo 18 del Estatuto (IMT, 1947, Vol. 1). La ley obligaba a los jueces a mantener una 'vista expeditiva' (*expeditious hearing*) de los cargos (p. 19) y a tomar medidas estrictas para evitar 'retrasos irrazonables' (*unreasonable delay*). Así, la interpretación simultánea se convirtió en un campo de batalla técnico: mientras el estatuto exigía celeridad, la mediación lingüística ofrecía a la defensa una herramienta imprevista para dilatar el tiempo y reflexionar sobre las respuestas.

3. LA INTERPRETACIÓN EN LA ACTUALIDAD

Lo que ocurrió en la Sala 606 de Núremberg no fue solo un evento histórico para el derecho; fue el motor que profesionalizó la interpretación de conferencias y la judicial. Lo que empezó como una solución de emergencia para que el juicio no se detuviera, ha terminado siendo una carrera reglada, con estándares técnicos y éticos que hoy se enseñan en las universidades. El éxito de un intérprete hoy no depende solo de su talento con los idiomas, sino de una formación y unos protocolos internacionales que aseguran que la comunicación sea exacta. En este capítulo vamos a analizar cómo se trabaja hoy, comparando las modalidades actuales con esos primeros pasos que se dieron en 1945.

3.1. Modalidades actuales: simultánea, consecutiva y judicial

En la práctica profesional de hoy, la interpretación se divide en varias modalidades según lo que necesite cada situación. Como explica Pöchhacker (2004, p. 19), esta separación se hizo oficial en los años 20, cuando la tecnología permitió diferenciar entre la interpretación consecutiva (traducir después del orador) y la simultánea (hacerlo a la vez). Este cambio fue el gran éxito de Núremberg y, de hecho, Baigorri-Jalón (2000, p. 270) lo resume bien con la frase «la función creó el órgano»: fue la presión por entenderse rápido en el tribunal lo que obligó a organizar el sistema profesional que usamos hoy.

Si miramos la interpretación consecutiva actual, ya no se trata solo de "traducir de memoria". Pöchhacker la describe como un "continuum" que cambia según lo largo que sea el discurso (2004, p. 19). A diferencia de los años 40, donde se confiaba mucho en la retentiva del intérprete, hoy se estudian técnicas específicas de toma de notas (*note-taking*). Estas notas no son una transcripción, sino símbolos que representan ideas y permiten reproducir discursos largos con una fidelidad total. Por otro lado, la interpretación judicial sigue manteniendo el principio de precisión que ya pedía el Artículo 25 del Estatuto de Núremberg. En los juicios actuales es muy normal usar la "interpretación a la vista", que es cuando el intérprete tiene que

leer y traducir al momento documentos escritos, como actas o sentencias, algo que exige una agilidad mental enorme.
La simultánea es la modalidad que más ha cambiado gracias a la tecnología. En Núremberg, los intérpretes trabajaban en cabinas de cristal abiertas por arriba, así que el ruido de la sala les entraba constantemente y les dificultaba concentrarse (Baigorri-Jalón, 2000, p. 275). Hoy, como señala Pöchhacker (2004, p. 20), lo normal es trabajar en cabinas insonorizadas y con sonido digital. Esto no es solo por comodidad; es una necesidad técnica para que el cerebro del intérprete pueda gestionar la carga cognitiva que supone escuchar y hablar a la vez sin distracciones externas.

Además de la cabina tradicional, existen otras formas de trabajar simultáneamente. Un ejemplo es el *chuchotage* o interpretación susurrada, donde el intérprete traduce al oído del cliente en voz baja. También existe la técnica "SimConsec", una mezcla moderna donde se usa una grabación para que el intérprete pueda traducir simultáneamente lo que acaba de oír de forma más precisa (Pöchhacker, 2004, p. 18). Por último, el sistema de *relay* o "lengua puente", que en Núremberg se usó por obligación para el ruso, es hoy fundamental en la ONU o la Unión Europea. Gracias al *relay*, un discurso en un idioma menos común puede pasar a un idioma "puente" y, desde ahí, el resto de intérpretes lo traducen a sus respectivas lenguas (Baigorri-Jalón, 2000, p. 270; Pöchhacker, 2004, p. 21).

3.2. Competencias del intérprete profesional

Si comparamos la época de Núremberg con la actualidad, vemos que las competencias exigidas a un intérprete han evolucionado desde la mera resistencia física hacia un perfil profundamente profesionalizado. Las competencias actuales son las propias de un oficio que se ha consolidado a través de la formación universitaria y los estudios académicos; son competencias mucho más desarrolladas por las que debe pasar el intérprete antes de poder entrar en cabina. En 1945, por el contrario, la principal preocupación era si el intérprete podía aguantar el ritmo. Como señala Baigorri-Jalón (2000, p. 253) citando a Birse, la interpretación simultánea se veía como un proceso "sumamente molesto" para el orador y agotador para el profesional, quien debía alcanzar velocidades de más de cien palabras por minuto, algo que casi iguala a una conversación normal.
En aquel entonces, la competencia se medía por la capacidad de no quedarse atrás para poder transmitir todo el discurso. Existían cifras contradictorias sobre qué ritmo era viable: mientras autores como los Bowen hablaban de 60 palabras por minuto (una velocidad de dictado), otros aseguran que se llegaba a las 130 o incluso 200 palabras cuando se leían textos ya traducidos (Baigorri-Jalón, 2000, p. 299). Lo más interesante es que no había una medición científica del esfuerzo; los intérpretes simplemente “aprendían sobre la marcha” y ganaban soltura conforme perdían el miedo escénico y automatizaban soluciones para términos difíciles.

Hoy en día, gracias a autores como Pöchhacker (2016, pp. 90-91), entendemos que lo que hacían esos pioneros era gestionar una "carga cognitiva" (cognitive load) altísima. Según el modelo de esfuerzos de Gile que recoge Pöchhacker, el intérprete debe repartir su capacidad mental en cuatro tareas simultáneas: escuchar y analizar (L), producir el discurso (P), gestionar la memoria (M) y coordinar todo el proceso (C). Para que la interpretación sea exitosa, la suma de estos esfuerzos no puede superar la capacidad total del cerebro.

Esta gestión de la energía mental es lo que garantiza la fidelidad (*fidelity*), que es la demanda principal de cualquier interpretación moderna. Como indica Pöchhacker (2016, pp. 135-136), ya no basta con transferir palabras como si fuera una "amplificación de sonido", sino que se debe transmitir "fiel y totalmente" las ideas del orador. En este sentido, la competencia actual se define bajo la norma del "portavoz honesto" (*honest spokesperson*): el intérprete debe reexpresar las ideas y la manera de expresarlas del orador original de la forma más exacta posible, evitando omisiones significativas (Harris, 1990, citado en Pöchhacker, 2016).

Finalmente, mientras que en Núremberg el control del ritmo dependía de factores externos —como el monitor que usaba una luz amarilla para pedir lentitud o una luz roja para detener al orador (Baigorri-Jalón, 2000, p. 300)—, el intérprete moderno cuenta con tácticas de prevención o *coping* para evitar el colapso mental. Ahora, una competencia clave es el "auto-monitoreo" constante. Esto implica que el profesional debe ser capaz de procesar la entrada de información mientras controla su propia producción de voz, ya que una entonación o prosodia inadecuada puede reducir drásticamente la comprensión de la audiencia (Pöchhacker, 2016, p. 140).

3.3. Requisitos en contextos jurídicos

La interpretación en el ámbito judicial actual se rige por requisitos de rigor y ética que, en gran medida, surgieron como respuesta a los problemas detectados en los juicios de la posguerra. El fundamento legal de esta práctica se establece en el Artículo 25 del Estatuto del Tribunal Militar Internacional (International Military Tribunal, 1947/2021), que reconoce el derecho del acusado a ser asistido por un intérprete. Este artículo no solo garantizaba la comunicación, sino que sentó las bases para considerar la interpretación como un derecho humano dentro de un proceso judicial justo.

Sin embargo, los requisitos actuales han tenido que evolucionar para evitar errores como los ocurridos en los inicios de la profesión. Como relata Baigorri-Jalón (2000, p. 286), en Núremberg se dieron casos donde la falta de distancia emocional o el exceso de recato afectaron a la traducción. Un ejemplo notable fue el de una intérprete que se bloqueó mentalmente al

pensar en su historia personal frente a los acusados; otro caso fue el de una profesional que se negaba a traducir palabras malsonantes del alemán al inglés, obligando a que un segundo intérprete interviniera solo para traducir los insultos.

Estos incidentes demuestran por qué hoy se exige lo que Mikkelson (2015, p. 567) denomina el "modelo de conducto" (*conduit model*). Según este requisito, el intérprete judicial debe actuar como un canal invisible y neutral, sin permitirse filtrar el lenguaje ni omitir términos, por muy ofensivos que resulten. La precisión actual exige que el mensaje llegue íntegro al juez o al jurado para que estos puedan valorar la prueba sin interferencias. En este sentido, un requisito crítico es no "limpiar" ni estructurar los testimonios que parecen fragmentados o confusos. Como advierte Mikkelson (2015, p. 372), un discurso aparentemente ilógico o contradictorio puede ser interpretado por el tribunal como una falta de credibilidad o una "mentira"; si el intérprete intentara dar coherencia a ese relato para ayudar al orador, estaría alterando una prueba fundamental para la evaluación del juez.

Asimismo, la profesionalización actual exige que los valores, creencias o normas del intérprete no interfieran en el contenido. Mikkelson recoge investigaciones donde se demuestra que los prejuicios personales pueden "filtrarse" de forma inconsciente, llegando incluso a cambiar el sentido de una declaración por completo (2015, p. 448). Se han documentado casos en los que, debido a los valores culturales del intérprete, consejos terapéuticos como "hablar de los problemas es bueno" se tradujeron como "es mejor callar para proteger a la familia", alterando totalmente la intención original.

En la actualidad, como señala Pöchhacker (2016), estos requisitos se han formalizado en códigos deontológicos que obligan a la imparcialidad y la fidelidad absoluta. Ya no se acepta que un intérprete "suavice" el discurso, lo reestructure o deje que sus propias emociones y valores influyan en la traducción, ya que su función es garantizar la igualdad de condiciones ante la ley. Estos estándares modernos aseguran que el derecho reconocido desde Núremberg se cumpla con una calidad técnica que evite cualquier distorsión de la justicia por motivos lingüísticos o personales.

4. INTELIGENCIA ARTIFICIAL APLICADA A LA INTERPRETACIÓN

4.1. Tecnologías actuales: reconocimiento de voz, traducción automática y LLM

La integración de la tecnología en la interpretación no es un fenómeno reciente, sino una evolución que comenzó con los sistemas electroacústicos de principios del siglo XX. Como señala Pöchhacker (2016, p. 182), desde la década de 1990 la convergencia de las telecomunicaciones y el procesamiento de datos digitales ha impactado profundamente la práctica profesional, pasando de herramientas que asisten al intérprete a sistemas que buscan replicar sus habilidades de forma autónoma.

4.1.1. Evolución del reconocimiento de voz (ASR)

El primer pilar de la tecnología actual es el Reconocimiento Automático de Voz (ASR). Tradicionalmente, este proceso se limitaba a convertir ondas sonoras en texto plano con un alto margen de error. Sin embargo, tecnologías de vanguardia como Whisper AI de OpenAI han transformado este campo. Este modelo utiliza redes neuronales avanzadas para analizar patrones de audio y contexto, logrando precisiones cercanas al 99% en más de 100 idiomas. A diferencia de los sistemas tradicionales, permite la transcripción instantánea, el manejo de múltiples hablantes y la resistencia al ruido ambiental, facilitando la base textual necesaria para cualquier proceso de traducción posterior (OpenAI, s.f.).

4.1.2. De los sistemas en cascada a los modelos "End-to-End"

Tradicionalmente, la interpretación automática se ha basado en sistemas en cascada. Según explica Pöchhacker (2016, p. 189), este modelo depende de tres componentes secuenciales: reconocimiento de voz (ASR), traducción automática (MT) y síntesis de voz (TTS). El principal inconveniente de esta arquitectura es la acumulación de errores de un paso a otro y la alta latencia (retraso).

La superación de este límite llega con los modelos "End-to-End" (extremo a extremo). Un hito en esta área es el sistema SeamlessM4T de Meta, un modelo unificado y multimodal que realiza traducciones de voz a voz sin depender de subsistemas separados. Este modelo mejora hasta en un 20% los estándares previos en traducción directa y reduce drásticamente la toxicidad y los sesgos, permitiendo una comunicación más fluida y segura en entornos globales (Meta AI, 2023).

4.1.3. Interpretación Asistida por Ordenador (CAI) y apoyo en cabina

Más allá de la automatización total, la tecnología ha transformado el rol del intérprete humano mediante las herramientas de CAI. Mientras que el equipamiento clásico se centraba en la transmisión del sonido (Pöchhacker, 2016, p. 183), las aplicaciones modernas se integran en el flujo de trabajo en tiempo real. Como indican Lu y Fantinuoli (2025), el uso de ASR permite extraer automáticamente nombres y cifras para mostrarlos al intérprete. Ejemplo de ello es BoothMate, una aplicación que optimiza la gestión terminológica y la preparación en cabina, permitiendo que el profesional consulte bases de datos complejas de forma instantánea para reducir su carga cognitiva (BoothMate, s.f.).

4.1.4. IA Generativa y plataformas de interpretación autónoma

El último estadio tecnológico es la irrupción de la IA Generativa y los LLM, que permiten gestionar matices lingüísticos y disfluencias (dudas, repeticiones) con una naturalidad antes inexistente. Plataformas como Wordly aprovechan estos motores patentados de IA para ofrecer interpretación remota y simultánea sin intérpretes humanos, permitiendo a los asistentes escuchar audio traducido o leer subtítulos en tiempo real (Wordly, s.f.). Este avance, según Lu y Fantinuoli (2025), representa un cambio de paradigma en la asequibilidad y escalabilidad de los eventos multilingües, permitiendo una colaboración global inmediata.

4.2. Impacto de la IA en la interpretación: beneficios y desafíos críticos

La implementación de la inteligencia artificial en la interpretación contemporánea ha generado lo que algunos expertos denominan la "paradoja de la automatización". Por un lado, sistemas como SeamlessM4T (Meta, 2023) y Whisper AI (OpenAI) ofrecen una capacidad de procesamiento masivo y una reducción drástica de la latencia, democratizando el acceso a la comunicación en entornos multilingües. Según Lu y Fantinuoli (2025), el modelo *human-in-the-loop* permite al intérprete liberarse de la carga de memorizar datos técnicos o cifras, delegando en la máquina la extracción terminológica para centrarse en la coherencia global del discurso.

Sin embargo, esta ventaja conlleva el riesgo de la "complacencia automatizada". El intérprete, bajo altos niveles de estrés, puede desarrollar una confianza excesiva en los resultados de la IA, lo que facilita que las "alucinaciones" del algoritmo —datos falsos generados con apariencia de veracidad— pasen desapercibidas. La IA tiende a priorizar la corrección gramatical sobre la precisión pragmática, lo que, en palabras de Lu y Fantinuoli (2025), puede derivar en una traducción literal que vacía de sentido el mensaje original, creando una falsa sensación de fluidez que oculta errores de fondo.

4.3. Limitaciones en el contexto legal: la alteración de la evidencia

En el ámbito judicial, la interpretación no es un mero intercambio de información, sino una pieza clave de la actividad probatoria. Aquí es donde la IA muestra sus mayores carencias. Según Mikkelson (2015, p. 341), los códigos de ética profesional prohíben estrictamente editar, resumir, borrar o añadir contenido; el intérprete debe conservar el registro, el estilo y el tono del orador. La IA, sin embargo, tiende a la "normalización": limpia el lenguaje confuso, corrige errores gramaticales del testigo y estandariza el discurso. Al hacer esto, la máquina altera la evidencia, pues impide que el juez evalúe la competencia, el estado mental o la veracidad de un declarante a través de sus propias vacilaciones o errores lingüísticos.

Además, Mikkelson (2015, p. 341) subraya que ante cualquier duda o impedimento para cumplir con la competencia exigida, el intérprete debe informar inmediatamente a la autoridad judicial. Una IA no tiene la capacidad de autoevaluarse ni de reconocer sus propias limitaciones; no puede "pedir aclaraciones" para evitar una mala interpretación, sino que simplemente ofrece el resultado estadísticamente más probable. Esta incapacidad de gestionar la duda convierte a la IA en un riesgo para la seguridad jurídica, ya que carece de la "sintonía fina" necesaria para detectar matices paralingüísticos como el sarcasmo o el miedo, elementos que definen la "calidad emocional" del encuentro judicial (Mikkelson, 2015, p. 454).

4.4. Marco ético y requisitos legales: la responsabilidad y la "caja negra"

El despliegue de la tecnología en la justicia se enfrenta hoy a una barrera ética fundamental: la rendición de cuentas (*accountability*). Mientras que el intérprete humano debe "mantener la apariencia de imparcialidad y neutralidad" y revelar cualquier conflicto de interés (Mikkelson, 2015, p. 341), los sistemas de IA operan como una "caja negra". La opacidad algorítmica impide que las partes en un juicio puedan confrontar o interrogar el origen de una traducción errónea, vulnerando el derecho fundamental a la defensa.

Esta peligrosidad es la que ha llevado a la Unión Europea en su Ley de IA (2024) a calificar el uso de estas tecnologías en la administración de justicia como de "alto riesgo". A diferencia de los profesionales de la era de Núremberg, que asumieron una responsabilidad moral y legal histórica sobre sus palabras, la IA carece de agencia ética. Como se analiza en la comparativa posterior, el papel del intérprete humano resulta insustituible no solo por su capacidad lingüística, sino por su obligación deontológica de actuar como garante de la justicia, corrigiendo errores sobre la marcha y asegurando que el registro y el tono del acusado se preserven íntegramente, sin las distorsiones propias de un procesamiento puramente estadístico.

5. EL DEBATE ENTRE LA AUTOMATIZACIÓN DIGITAL Y EL JUICIO HUMANO EN EL ÁMBITO PENAL

El recorrido teórico desde los orígenes de la interpretación simultánea en Núremberg hasta el escenario tecnológico actual permite observar que la profesión no parece dirigirse hacia una sustitución total de los intérpretes por máquinas, sino más bien hacia una transformación del modo de trabajo y del control del proceso judicial. Si se compara el contexto de la posguerra con el desarrollo actual de la inteligencia artificial, puede observarse que su aplicación en los tribunales todavía presenta importantes limitaciones, especialmente porque un entorno tan estricto como el derecho procesal penal exige unas garantías de seguridad que la tecnología actual aún no puede garantizar.

Como advierten investigadores y miembros del Comité Técnico de la Asociación Internacional de Intérpretes de Conferencias (AIIC), existe un sesgo inherente al utilizar el término "inteligencia" en el contexto informático, ya que la capacidad de las máquinas es infinitamente más estrecha que la complejidad del pensamiento humano (Pereira Ginet-Jaquemet, 2023). En realidad, los modelos de lenguaje actuales no realizan una interpretación en el sentido humano del término, sino un procesamiento del lenguaje basado en cálculos matemáticos, almacenamiento de datos y patrones estadísticos, sin incorporar elementos esenciales de la comunicación humana. Por ello, en el ámbito académico y legal es más preciso hablar de soluciones de traducción oral automática guiadas por IA que de interpretación automatizada. Esta visión prudente también es compartida por las principales organizaciones del sector, como la propia AIIC a través de su Grupo de Trabajo sobre Inteligencia Artificial (*Artificial Intelligence Task Force* o AITF), creado específicamente para evaluar el impacto de estas tecnologías y preservar los matices humanos que las máquinas son incapaces de replicar (Pereira Ginet-Jaquemet, 2023).

5.1. El "efecto cascada" y los riesgos cognitivos en el tribunal

Para comprender por qué estas soluciones automáticas son inviables de forma autónoma en un tribunal, los análisis especializados explican que la Traducción Automática del Habla (AST) es un proceso ciego de tres pasos secuenciales que comprende el reconocimiento automático del habla para pasar de voz a texto, la traducción automática de dicho texto y, finalmente, la conversión de texto a voz mediante síntesis generativa (Goldsmith y Díaz, 2024). El riesgo en el ámbito judicial es que los errores se acumulan en un "efecto cascada", lo que significa que, si el software comete un fallo al transcribir un nombre propio o un término especializado en el primer paso, arrastra y multiplica ese error semántico en la traducción y en la locución final de forma inevitable (Goldsmith y Díaz, 2024). Esta falta de precisión se debe a que los modelos

de aprendizaje automático dependen por completo de la calidad de los datos masivos con los que son entrenados, considerados actualmente el nuevo oro negro de la tecnología (Pereira Ginet-Jaquemet, 2023). Sin embargo, datos actuales distan mucho de ser cien por cien exactos o reales, acumulando sesgos, distorsiones y lo que la literatura especializada denomina basura informática o *data litter*, lo que provoca rendimientos deficientes o predicciones contradictorias ante discursos judiciales complejos (Pereira Ginet-Jaquemet, 2023).

Ante las limitaciones de la automatización pura, la neurociencia cognitiva aplicada a la interpretación ha comenzado a evaluar cómo interactúa el cerebro humano cuando se introducen herramientas de asistencia tecnológica en la cabina. En este ámbito, un pionero estudio sociocognitivo desarrollado por la Universidad de Gante en colaboración directa con la Dirección General de Interpretación de la Comisión Europea (DG SCIC) ha aportado las primeras evidencias científicas biométricas utilizando técnicas avanzadas de rastreo ocular (*eye tracking*) y electroencefalografía (EEG) continua (Pleșca et al., 2024). Al analizar el comportamiento de profesionales expuestos al sistema de asistencia en tiempo real *The Artificial BoothMate*, los resultados demostraron la existencia de un gran "efecto de sustitución" en el procesamiento visual (Pleșca et al., 2024). Cuando la herramienta tecnológica permanece apagada, los intérpretes se ven obligados a sobrecargar su atención consultando constantemente apuntes en papel y glosarios tradicionales, llegando a dedicar un 31.67% de su tiempo a la lectura de glosarios o documentos impresos y un 12.52% a las diapositivas proyectadas en la sala (Pleșca et al., 2024).

Sin embargo, al activar el dispositivo de soporte con transcripción automática y extracción terminológica en directo, la fijación visual sobre los documentos en papel decae drásticamente al 12.62% y la atención sobre las diapositivas disminuye a la mitad (Pleșca et al., 2024). Lejos de provocar una sobrecarga mental alienante, el uso de la pantalla de asistencia absorbe únicamente el 18.53% del tiempo visual y desencadena un fenómeno de liberación cognitiva, permitiendo al intérprete dedicar la mayor parte de su atención (el 49.18%) a monitorizar de forma directa la interacción humana, los gestos y los turnos de palabra de los participantes en el tribunal (Pleșca et al., 2024). Los datos del rastreo ocular demuestran que, aunque los intérpretes miran fijamente la pantalla durante más tiempo que los apuntes en papel, esto no significa que les cueste más esfuerzo mental. Simplemente se debe a que la pantalla es muy dinámica y capta su atención porque se actualiza constantemente con nuevas líneas de texto y números (Pleșca et al., 2024).

A estos problemas de lenguaje se suma el riesgo de que el intérprete desarrolle demasiada confianza en la tecnología automática en momentos de mucha presión (Lu y Fantinuoli, 2024). Cuando un profesional se encuentra sometido a una presión extrema, existe el riesgo neurocognitivo de reducir la atención y validar ciegamente las propuestas automáticas de la

pantalla para aliviar la fatiga visual, lo que podría provocar el incremento de errores graves en el registro judicial de las declaraciones (Lu y Fantinuoli, 2024).

5.2. El criterio institucional de la AIIC y los escenarios de alto riesgo judicial

Ante los fallos técnicos de la traducción automática y el peligro de que se introduzcan errores graves en un juicio, la propia AIIC, a través de su grupo de investigación tecnológica (*AI Workstream*), publicó oficialmente el 16 de enero de 2025 una herramienta de evaluación denominada el Árbol de Decisión de la IA o *AI Decision Tree* (Goldsmith, 2025). Este recurso de orientación institucional deja claro que la traducción automática solo puede llegar a ser una opción si se cumplen unas condiciones extremadamente de laboratorio que resultan prácticamente imposibles para un tribunal, tales como que la interpretación no sea una parte crítica para el éxito del evento, que los discursos estén completamente escritos o guionizados de antemano y que el mensaje esté totalmente desprovisto de emociones, humor, dobles sentidos o referencias culturales (Goldsmith, 2025). En la práctica, estas condiciones resultan muy difíciles de aplicar en un tribunal penal, ya que los interrogatorios son impredecibles, los testigos improvisan constantemente y las declaraciones están cargadas de la crudeza emocional y los detalles culturales de los delitos juzgados (Goldsmith, 2025).

Desde la perspectiva defendida por los comités de expertos de la AIIC, por más avanzados que sean los sistemas de computación a la hora de superar el test de Turing o el juego de la imitación de datos, seguirá habiendo una diferencia difícil de salvar entre el intérprete humano y la máquina, porque, a diferencia del algoritmo, el ser humano posee la capacidad real de comprender el significado profundo de lo que está diciendo (AIIC, 2020). Mientras que la máquina ejecuta un procesamiento ciego de códigos y probabilidades estadísticas, el profesional aplica la intuición, la empatía, el conocimiento cultural y la etiqueta social para descifrar la intención del orador y el trasfondo del testimonio (AIIC, 2020). Por estas razones, la postura oficial de la institución sostiene que en escenarios de alta responsabilidad o alto riesgo (*high-stakes events*), donde el coste del error es inasumible, la interpretación humana sigue siendo la única opción fiable y legalmente viable para garantizar los derechos procesales (AIIC, 2020).

Esta necesidad de supervisión humana está alineada con las normativas recientes de la Unión Europea. De hecho, la Comisión Europea ha clasificado el uso de la Inteligencia Artificial en los tribunales como una actividad de "alto riesgo". Por este motivo, la ley exige que cualquier herramienta tecnológica en un juicio esté sometida a un control humano estricto, garantizando así que no se vulneren los derechos fundamentales de los ciudadanos (AI Act Service Desk, 2024).

5.3. La IA como asistente de cabina y el control del intérprete

Aun así, también es importante reconocer que, a pesar de que la IA carece de validez ética y legal para actuar de forma autónoma dentro de la sala de vistas, se ha convertido en una herramienta de apoyo especialmente útil para la fase previa al juicio o como soporte complementario en cabina. Los análisis neurofisiológicos derivados de la monitorización cerebral por EEG de la Universidad de Gante confirman que la integración de herramientas de asistencia tecnológica no ejerce un impacto negativo ni perjudicial sobre la capacidad cognitiva del profesional (Pleșca et al., 2024). Al estudiar el comportamiento de las ondas cerebrales a través del Índice de Carga Cognitiva (*CLI*), que calcula la relación entre la potencia theta frontal *(indicador del esfuerzo de atención sostenida)* y alpha parietal *(vinculado a la memoria y el procesamiento lingüístico)*, los neurocientíficos constataron que el soporte tecnológico apenas reduce el desgaste cognitivo real que sufre el cerebro (Pleșca et al., 2024).

Según estos estudios, el principal origen del agotamiento cognitivo y del incremento de las ondas theta en el cerebro del intérprete no radica en la interacción con los dispositivos digitales, sino en el factor temporal y el avance del tiempo cronológico del turno de trabajo, lo que consolida la tesis de que la fatiga mental es un proceso acumulativo derivado de la propia actividad interpretativa (Pleșca et al., 2024). Asimismo, los análisis de la voz demuestran que la presencia de la IA no afecta al tono ni a la calidad del habla del intérprete, manteniendo intacto su rendimiento profesional (Pleșca et al., 2024). En conjunto, los datos sugieren que estas tecnologías pueden funcionar como una herramienta de apoyo al intérprete. Lejos de saturar al profesional, le ayudan a resolver dudas de vocabulario muy específicas en momentos de gran acumulación técnica (Pleșca et al., 2024).

A través de herramientas avanzadas de gestión terminológica diseñadas específicamente para el sector, como *BoothMate*, algunos intérpretes actuales han dejado atrás los formatos estáticos tradicionales de Word o Excel para dar el salto a los glosarios multimedia (Goldsmith, 2024d). La investigación en este campo confirma que la integración de la IA en la gestión de glosarios mejora la preparación previa del intérprete. El uso de recursos visuales —como diagramas o mapas— vinculados a conceptos jurídicos difíciles facilita el recuerdo en cabina, puesto que el cerebro humano procesa y retiene las imágenes con mayor eficacia que los textos complejos (Goldsmith, 2024d). Del mismo modo, el uso de aplicaciones que recopilan archivos de audio con pronunciaciones nativas permite asegurar la precisión de la dicción de nombres propios extranjeros o tecnicismos médicos antes del discurso oral, reduciendo la carga mental del profesional (Goldsmith, 2024d).

A esta optimización terminológica se suma la aparición de entornos cerrados de inteligencia artificial como *NotebookLM*, que se han convertido en herramientas clave para combatir la saturación de información durante la fase de documentación previa (Goldsmith, 2025). La gran

ventaja de este tipo de tecnología aplicada a la interpretación es que, a diferencia de los modelos generales de inteligencia artificial que extraen datos de todo internet y tienden a generar errores de precisión constantes, estos sistemas trabajan de forma exclusiva con los archivos y fuentes de confianza que el propio profesional introduce. Ello permite al intérprete interrogar simultáneamente decenas de documentos judiciales, extraer resúmenes en forma de guías de estudio y localizar datos de forma inmediata en situaciones de alta presión. En este sentido, el impacto real de la inteligencia artificial no parece ser sustituir al intérprete, sino transformar la manera en que este busca información y prepara el caso antes de entrar al tribunal (Goldsmith, 2025).

En conclusión, el marco teórico analizado permite sostener que el factor humano sigue siendo indispensable en los tribunales por tres razones que, al menos actualmente, la IA todavía no puede reproducir plenamente: en primer lugar, la adaptabilidad pragmática para entender la ironía o los titubeos de un testigo; en segundo lugar, la sensibilidad lingüística necesaria para procesar testimonios traumáticos; y, por encima de todo, la responsabilidad ética y legal. Mientras que en 1945 los pioneros de Núremberg disponían de la autoridad del interruptor a través del sistema de luces de IBM para detener el juicio ante la pérdida de sentido (Gaiba, 1998), un algoritmo carece de criterio jurídico propio y sigue ofreciendo resultados sin ser capaz de detectar por sí mismo cuándo se ha producido un error relevante. En un contexto penal, donde un fallo de traducción puede comprometer derechos fundamentales e incluso afectar a la libertad de una persona inocente, el estado actual del desarrollo tecnológico sigue planteando riesgos significativos para una delegación autónoma de la interpretación judicial en sistemas de inteligencia artificial, especialmente en lo relativo a la preservación de las garantías procesales. Estas consideraciones teóricas justifican y dan sentido a la parte práctica de este trabajo, orientada a evaluar empíricamente, mediante fragmentos de juicios reales, las limitaciones que aún separan la traducción automática de las exigencias éticas, pragmáticas y lingüísticas propias de la interpretación judicial humana.

6. ANÁLISIS PRÁCTICO DE LA INTERPRETACIÓN AUTOMÁTICA EN SEDE JUDICIAL

Tras exponer el debate sobre la automatización en el ámbito jurídico, el trabajo analiza el funcionamiento real de estas herramientas. Este apartado se centra en un caso judicial concreto para analizar cómo responde la inteligencia artificial ante un testimonio marcado por la complejidad lingüística y emocional propia de un acusado. El objetivo es contrastar la teoría con la práctica y analizar si la traducción automática logra mantener la fidelidad necesaria o si sus limitaciones afectan a la precisión que puede ofrecer un intérprete humano. Con este análisis, se busca responder a si es ético utilizar estos sistemas en casos judiciales sensibles y

hasta qué punto la inteligencia artificial puede asumir las funciones del intérprete sin sacrificar la exactitud lingüística y los estándares éticos exigidos.

6.1. Metodología de la investigación

Para el desarrollo de este estudio, se ha seleccionado la declaración del acusado en el juicio de Kenosha (2020). Este caso presenta un testimonio de alta complejidad, ya que la declaración de Kyle Rittenhouse destaca por presentar momentos de gran tensión emocional, incluyendo episodios de llanto visibles, lo que permite analizar cómo la máquina procesa las vacilaciones, pausas, silencios y episodios de llanto dentro del testimonio.

Se ha elegido DeepL Voice para este análisis por tratarse de una herramienta de traducción automática ampliamente accesible y representativa de las soluciones de IA de consumo actualmente disponibles. Si bien existen plataformas de interpretación simultánea remota (RSI) orientadas al ámbito profesional —como *Kudo* o *Wordly*—, estas integran motores de procesamiento de lenguaje natural y arquitecturas de traducción automática que comparten fundamentos tecnológicos similares a los de las soluciones de uso general. A diferencia de las plataformas RSI, diseñadas para ser utilizadas por intérpretes profesionales en entornos controlados, el foco de este análisis recae en la tecnología de traducción automática subyacente. Más allá de la interfaz o del entorno de uso, esta tecnología sigue enfrentándose a limitaciones similares cuando se aplica a un contexto judicial penal de alta exigencia comunicativa y probatoria.

El procedimiento seguido para esta parte práctica se ha dividido en tres fases. En primer lugar, se procesó el audio del fragmento del discurso original mediante interpretación en tiempo real con la herramienta DeepL Voice. Esta función, disponible únicamente en la versión de pago (Pro), presenta limitaciones técnicas relevantes para un contexto de interpretación simultánea en tiempo real. Tras probar las distintas opciones que ofrece la aplicación, se observó que herramientas integradas con plataformas como Zoom o Microsoft Teams no resultaban adecuadas para este estudio, ya que la salida de la traducción se ofrecía exclusivamente en formato textual, sin reproducción oral simultánea, lo que impedía una comparación adecuada con las condiciones propias de la interpretación judicial.

Por este motivo, se optó por la función de conversación grupal de DeepL Voice, que permite a cada participante seleccionar un idioma de destino y recibir la interpretación de forma oral y prácticamente simultánea. Para llevar a cabo la prueba, fue necesario utilizar dos dispositivos: uno configurado como emisor, con el idioma original en inglés (Estados Unidos), y otro como receptor, configurado en español. Conviene señalar que, en esta modalidad, DeepL únicamente permite seleccionar la variante de español latinoamericano.

Debido a las limitaciones de funcionamiento de la propia plataforma, el audio original tuvo que reproducirse desde una fuente externa para que el dispositivo emisor pudiera captarlo correctamente. De este modo, fue posible registrar tanto el audio original como la interpretación generada por la IA desde el dispositivo receptor. Para facilitar la verificación de los resultados y garantizar una mayor transparencia del procedimiento, se ha incluido un código QR junto con un enlace directo al vídeo completo de la prueba realizada en el anexo.
Finalmente, se procedió a la transcripción del output generado por el sistema y a la elaboración de una tabla comparativa entre el discurso original, la versión producida por DeepL Voice y una propuesta de interpretación humana de referencia. Este análisis permite evaluar el rendimiento de la herramienta, especialmente en relación con la precisión lingüística, la gestión de matices pragmáticos y emocionales, y el cumplimiento de los estándares éticos exigidos en el ámbito judicial

6.2. Análisis de resultados: de la técnica a la ética procesal

Para evaluar la fiabilidad de *DeepL Voice*, se ha analizado un fragmento del interrogatorio a Kyle Rittenhouse. En este segmento, el acusado relata los momentos previos a los disparos, describiendo cómo se encontraba en una situación de extrema tensión mientras intentaba apagar un incendio. El discurso es narrativo y complejo, con detalles sobre ubicaciones, personas involucradas y la percepción del peligro por parte del acusado.

A continuación, se presenta la comparativa entre la intervención original, la interpretación generada por DeepL Voice y la propuesta de referencia, que servirá de base para el análisis posterior (Tabla 1):

Interviniente	Transcripción original (EN)	DeepL Voice	Interpretación humana de referencia (ES)
Abogado	"Describe your approach to Car Source number three."	"Describe tu enfoque respecto a la fuente de datos número tres."	"Describa qué ocurrió cuando se acercó al aparcamiento de Car Source número tres."
Acusado (Kyle Rittenhouse)	"As I'm walking down Sheridan Road, I hear somebody scream, 'Burn in hell.'"	"Mientras camino por Sheridan Road, oigo a alguien gritar «¡Arde en el infierno!»"	"Mientras caminaba por Sheridan Road, oí a alguien gritar: '¡Arde en el infierno!'"
Acusado	"And I reply with, 'Friendly, friendly, friendly.' To let them know, hey, I'm just here to help."	"Y yo respondí con amabilidad, amabilidad, amabilidad para hacerles saber: «Oye, solo estoy aquí para ayudar»."	"Y yo respondí: 'Tranquilos, tranquilos, tranquilos', para hacerles saber: 'Oye, solo estoy aquí para ayudar'."
Acusado	"I don't want any problems. I just want to	"Solo es que no quiero problemas. Solo quiero	"No quería problemas. Solo quería apagar

	put out the fires if there are any."	apagar los incendios, si es que hay alguno."	cualquier incendio, si lo había."
Acusado	"I continue walking and then I notice the Duramax, I notice a flame in the back seat of the Duramax…"	"Sigo caminando y entonces me doy cuenta del Duramax. Me fijé en una llama en el asiento trasero del Duramax…"	"Seguí caminando y vi el Duramax; vi una llama en el asiento trasero…"
Acusado	"As I'm stepping forward, I believe his name is now Joshua Ziminski, he steps towards me with a pistol in his hand."	"Y mientras avanza, creo que ahora se llama Joshua Zaminski. Él se acercó a mí con una pistola en la mano."	"Mientras avanzaba, creo que su nombre es Joshua Ziminski, se acercó hacia mí con una pistola en la mano."
Acusado	"As I'm walking towards to put out the fire, I drop the fire extinguisher and I take a step back."	"Mientras caminaba, mientras caminaba hacia allí para apagar el fuego, se me cayó el extintor y di un paso atrás."	"Mientras me acercaba para apagar el fuego, solté el extintor y di un paso atrás."
Abogado	"When you step back from Mr. Ziminski, what's your plan?"	"Cuando das un paso atrás del Sr. Zeminski, ¿cuál es su plan?"	"Cuando se alejó del Sr. Ziminski, ¿cuál era su intención?"
Acusado	"My plan is to get out of that situation and go back north down Sheridan Road to where the Car Source lot number two was."	"Mi plan es salir de esa situación y volver al norte, a Sheridan Road, donde está el aparcamiento de coches número dos."	"Mi intención era salir de esa situación y regresar hacia el norte por Sheridan Road, donde se encontraba el aparcamiento de Car Source número dos."
Abogado	"And did you get back, were you able to go in a northerly direction?"	"¿Y consiguió volver? ¿Pudiste dirigirte hacia el norte?"	"¿Pudo regresar? ¿Consiguió dirigirse hacia el norte?"
Acusado	"I wasn't."	"No lo hice."	"No pude."
Abogado	"Describe what happens."	"Describa lo que ocurrió."	"Describa qué ocurrió después."
Acusado	"Once I take that step back, I look over my shoulder and Mr. Rosenbaum was now running from my right side…"	"Una vez que di ese paso atrás. Miro por encima del hombro y… Sr. Rosenbaum, Sr. Rosenbaum ahora corría por mi lado derecho."	"Después de dar ese paso atrás, miré por encima del hombro y el Sr. Rosenbaum venía corriendo desde mi lado derecho."
Acusado	"I was cornered from in front of me with Mr. Ziminski and there were three people right there."	"Y yo estaba acorralado por delante con el Sr. Zeminski. Y había…había gente justo ahí. Eso es lo que yo... Eso es lo que traje."	"Estaba acorralado, con el Sr. Ziminski delante de mí, y había tres personas justo allí. *El acusado llora y no puede continuar*"
Abogado	"Take a deep breath, Kyle."	—	"Respire hondo, Kyle."
Juez	"We're going to just take a… Time for our break anyway. You can just relax for a minute, sir."	"Vamos a tomarnos un momento para el descanso, de todos	"Vamos a hacer una pausa. De todos modos, ya es hora del descanso.

		modos, puedes relajarte un momento, señor"	Puede relajarse un momento, señor."
Juez	"We're going to take a break about 10 minutes and please don't talk about the case during the break, read, watch and listen to any comment on the trial."	"Vamos a hacer una pausa de unos 10 minutos, y por favor, no hable del caso durante la pausa. Lea, vea o escuche cualquier comentario."	"Vamos a hacer una pausa de unos diez minutos y, por favor, no hable del caso durante la pausa ni lea, vea o escuche comentarios sobre el juicio."

Tabla 1: ANÁLISIS COMPARATIVO DEL DISCURSO ORIGINAL, SALIDA DE DEEPL VOICE Y REFERENCIA HUMANA.

6.2.1. Análisis lingüístico

Desde una perspectiva lingüística, los resultados muestran que *DeepL Voice* logra transmitir el sentido general del discurso en gran parte del fragmento. No obstante, el análisis detallado revela errores relevantes de precisión terminológica, interpretación contextual y adecuación lingüística, especialmente problemáticos en un entorno judicial, donde pequeños matices pueden tener consecuencias significativas sobre la comprensión del testimonio.

Uno de los errores más evidentes aparece al inicio de la intervención, cuando la expresión *"Car Source number three"* es traducida como *"fuente de datos número tres"*. En este caso, la herramienta interpreta el término *source* de manera literal y no identifica que *Car Source* constituye el nombre propio de un lugar relacionado con los hechos del caso. La consecuencia de este error no es únicamente terminológica, sino también contextual, ya que altera la referencia a un espacio importante dentro de la narración. Frente a ello, una interpretación humana como *"aparcamiento de Car Source número tres"* mantiene el referente original y sitúa correctamente el escenario de los acontecimientos. Este ejemplo refleja una de las principales limitaciones de la IA en el reconocimiento contextual de nombres propios y elementos contextuales ajenos al significado puramente lingüístico.

Otro caso significativo es la interpretación de *"Friendly, friendly, friendly"*, traducida por *DeepL Voice* como *"amabilidad, amabilidad, amabilidad"*. Aunque desde un punto de vista léxico ambas expresiones presentan una relación semántica, la herramienta no logra captar el valor pragmático de la intervención. En este contexto, el acusado pretende comunicar a varias personas que no supone una amenaza y que sus intenciones son pacíficas. Por ello, una reformulación como *"tranquilos, tranquilos, tranquilos"* reproduce de forma más adecuada la intención comunicativa y refleja mejor la espontaneidad de la situación. Este caso muestra cómo la IA tiende a priorizar equivalencias léxicas frente a interpretaciones contextualizadas del significado. Además, este fragmento también pone de manifiesto ciertas dificultades

relacionadas con la naturalidad lingüística del discurso producido por la IA. Expresiones como "Solo es que no quiero problemas. Solo quiero apagar los incendios, si es que hay alguno" evidencian una tendencia a reproducir estructuras excesivamente cercanas al inglés y poco naturales en español oral espontáneo. La reiteración de "solo", junto con la construcción sintáctica empleada, genera un discurso menos fluido que formulaciones más naturales como "No quería problemas. Solo quería apagar cualquier incendio, si lo había". Del mismo modo, la expresión "Sigo caminando y entonces me doy cuenta del Duramax" evidencia una traducción demasiado literal del verbo *notice*, resultando menos idiomática en este contexto narrativo que opciones como "vi el Duramax" o "me fijé en el Duramax". Además, se observan ciertas dificultades de cohesión discursiva, ya que la repetición inmediata del referente en "Me fijé en una llama en el asiento trasero del Duramax" genera una narración más rígida y redundante. Algo similar ocurre con la repetición de estructuras próximas como "mientras caminaba" y "mientras caminaba hacia allí", que aproxima excesivamente la interpretación a la formulación original y reduce su naturalidad en español.

También se observaron algunas pérdidas de precisión en el significado que, en un contexto judicial, podrían resultar problemáticas desde el punto de vista probatorio. Un ejemplo relevante se encuentra en la frase *"there were three people right there"*, traducida como *"había gente justo ahí"*. La omisión del número concreto modifica el contenido del testimonio, ya que elimina un dato específico transmitido por el acusado sobre el número de personas presentes. En un procedimiento judicial, donde detalles aparentemente menores pueden resultar relevantes para la reconstrucción de los hechos, estas simplificaciones podrían afectar a la fidelidad de la declaración.

Asimismo, se identificaron problemas relacionados con la adecuación lingüística. Expresiones como *"I wasn't"* fueron interpretadas como *"No lo hice"*, una formulación poco natural y semánticamente incorrecta en el contexto, frente a una opción más adecuada como *"No pude"*. Del mismo modo, se observaron dificultades en el reconocimiento fonético de nombres propios, dando lugar a variaciones como *"Joshua Zaminski"* o *"Zeminski"* en lugar de *Ziminski*. Aunque estos errores no impiden completamente la comprensión global del mensaje, sí pueden generar imprecisiones relevantes en contextos judiciales, donde es esencial identificar correctamente a las personas. Estos errores ilustran cómo la herramienta prioriza la fluidez sobre la fidelidad probatoria, un aspecto especialmente delicado en el ámbito penal, donde la precisión es indispensable. Como advierte Pereira Ginet-Jaquemet (2023), la dependencia de la IA de modelos basados en cálculos probabilísticos, sumada a la calidad variable de los datos de entrenamiento, dificulta que estas soluciones alcancen la precisión necesaria para una interpretación fiable en contextos complejos.

6.2.2. Rendimiento técnico *DeepL Voice*

Desde el punto de vista técnico, la herramienta mostró un rendimiento relativamente fluido durante gran parte del fragmento analizado. A diferencia de otras alternativas exploradas previamente, la herramienta consiguió mantener una continuidad bastante estable del discurso y evitó silencios excesivamente largos en la mayor parte de la prueba. La interpretación se produjo con una latencia relativamente reducida, permitiendo seguir el hilo general de la intervención sin interrupciones constantes.

No obstante, sí se detectaron algunos momentos de silencio entre el discurso original y la salida oral traducida. Aunque estas pausas no fueron constantes ni llegaron a impedir completamente el seguimiento del mensaje, en determinados momentos podían resultar lo suficientemente prolongadas como para afectar a la sensación de inmediatez que exige una interpretación simultánea en un contexto judicial. Estas situaciones aparecieron con mayor frecuencia en fragmentos emocionalmente intensos o sintácticamente más complejos.
Otro aspecto especialmente relevante fue la segmentación irregular del discurso. En diversas ocasiones, *DeepL Voice* parecía interpretar que el hablante había finalizado una idea cuando todavía no había terminado de expresarse, cerrando frases antes de tiempo y realizando pausas poco naturales. Como consecuencia, en algunos momentos presentaba una entonación artificial o extraña, generando la sensación de que ciertas frases quedaban "cerradas" para posteriormente retomarse. Aunque este fenómeno no imposibilita la comprensión global del mensaje, sí afecta a la fluidez y naturalidad del discurso interpretado, alejándose del ritmo continuo que se espera de una interpretación profesional.

Como señalan Lu y Fantinuoli (2025), la naturaleza de la interpretación, que exige una cognición situada y un procesamiento en tiempo real, representa un desafío superior para los sistemas de IA frente a la traducción escrita, especialmente por la complejidad del lenguaje espontáneo. En este sentido, los cortes abruptos y las pausas extrañas de *DeepL Voice* no son solo un problema de calidad de sonido. Estos resultados sugieren que la IA todavía presenta dificultades para gestionar el ritmo y la estructura natural del habla humana durante la interpretación en tiempo real.

6.2.3. Dimensión emocional y ética procesal

Uno de los aspectos más relevantes del fragmento analizado es la elevada carga emocional del testimonio. A medida que avanza el relato, Kyle Rittenhouse muestra signos visibles de nerviosismo hasta terminar rompiéndose emocionalmente y siendo incapaz de continuar hablando con normalidad. Es precisamente en este momento donde *DeepL Voice* parece encontrar mayores dificultades interpretativas. La herramienta no interpreta el quiebre emocional del acusado como una interrupción de la declaración, sino que continúa intentando

producir lenguaje incluso cuando el audio deja de ser inteligible. Esto da lugar a frases incoherentes como *"Eso es lo que yo… Eso es lo que traje"*. Estas dificultades resultan especialmente relevantes si se tiene en cuenta que el intérprete judicial no solo debe trasladar el significado de las palabras, sino también preservar el registro, el estilo y el tono del hablante (Lee, 2015, p. 223). Precisamente en estos aspectos, especialmente durante los momentos de mayor carga emocional del acusado, *DeepL Voice* parece presentar mayores limitaciones.

Otro aspecto interesante observado durante el fragmento es la forma en que *DeepL Voice* gestiona las vacilaciones y repeticiones del hablante en momentos de tensión emocional. Cuando Kyle Rittenhouse comienza a mostrarse más nervioso y su discurso se vuelve menos fluido, aparecen repeticiones propias del habla espontánea, como *"Mr. Rosenbaum, Mr. Rosenbaum…"* o *"there were… there were three people right there"*. En lugar de omitir estas reiteraciones o reformular el discurso para hacerlo más fluido, la IA las reproduce de forma literal en la lengua meta (*"Sr. Rosenbaum, Sr. Rosenbaum…"* o *"había… había gente justo ahí"*). Este comportamiento muestra una cierta capacidad de la herramienta para conservar rasgos propios del habla espontánea, incluidas las vacilaciones del orador. No obstante, al trasladarlas de forma tan literal, el resultado puede sonar menos natural en español y generar un discurso más entrecortado y difícil de comprender. Esta cuestión resulta especialmente relevante si se tiene en cuenta que el contenido, la coherencia y el estilo del discurso de un testigo pueden influir en la percepción de credibilidad y fiabilidad por parte del tribunal (Lee, 2015, p. 232). En este sentido, cualquier alteración en el tono, las vacilaciones o la forma de expresar el testimonio puede tener implicaciones que van más allá de la mera traducción del contenido literal.

Además, conviene recordar que, tal y como señala Lee (2015, p. 227), el derecho a un juicio justo no puede separarse de la calidad de la interpretación, y la mera provisión de un servicio interpretativo no garantiza por sí sola la equidad procesal. En conjunto, aunque *DeepL Voice* consigue transmitir gran parte del contenido general del discurso, las limitaciones observadas en precisión terminológica, naturalidad, fidelidad contextual y gestión de situaciones emocionalmente complejas plantean dudas sobre su capacidad para garantizar una representación plenamente fiel del testimonio y satisfacer los estándares de precisión exigidos en el ámbito judicial.

7. CONCLUSIONES

El recorrido desarrollado a lo largo de este trabajo, desde los Juicios de Núremberg hasta las actuales herramientas de inteligencia artificial aplicadas a la traducción oral, permite observar que, aunque la tecnología ha transformado profundamente los medios de interpretación, el problema fundamental que dio origen a la interpretación judicial moderna continúa siendo esencialmente el mismo: garantizar que una persona pueda comprender y ser comprendida en un proceso judicial sin que la barrera lingüística comprometa la justicia. Por ello, el análisis histórico realizado no constituye únicamente un marco contextual, sino también un punto de partida para comprender los límites y posibilidades de la inteligencia artificial en la interpretación judicial contemporánea (Gaiba, 1998; Lee, 2015, pp. 224–227).

Los Juicios de Núremberg supusieron no solo el nacimiento de la interpretación simultánea moderna, sino también la consolidación de un principio que continúa vigente: la necesidad de asegurar una mediación lingüística precisa en contextos donde las consecuencias procesales pueden resultar irreversibles. La complejidad de aquellos procesos obligó a desarrollar mecanismos de control interpretativo, entre ellos el conocido sistema de luces de IBM, que permitía interrumpir la intervención cuando se producía una pérdida de sentido o un problema de comprensión (Gaiba, 1998). Más allá de su utilidad técnica, este sistema marcó un principio fundamental para la interpretación judicial actual. Aquella herramienta demostró que la fidelidad en el mensaje no es un simple detalle, sino la base para garantizar un juicio justo. Cuando las consecuencias en un tribunal son definitivas, la precisión del intérprete se convierte en una garantía obligatoria que asegura la justicia del proceso.

Desde esta perspectiva histórica, el interés creciente por la inteligencia artificial aplicada a la traducción oral nos hace plantearnos hasta qué punto los sistemas actuales pueden asumir funciones tradicionalmente desempeñadas por intérpretes humanos sin comprometer la exactitud lingüística ni los estándares éticos exigidos en el ámbito judicial. A primera vista, herramientas como *DeepL Voice* parecen representar un avance significativo en términos de accesibilidad, rapidez y reducción de la latencia, especialmente en un contexto cada vez más digitalizado. Sin embargo, tal y como advierten investigadores y miembros del Comité Técnico de la AIIC, el propio concepto de "inteligencia" aplicado a estos sistemas puede resultar engañoso, dado que la capacidad de las máquinas continúa siendo mucho más limitada que la complejidad del razonamiento humano. Más que interpretar el lenguaje en un sentido humano, estos sistemas funcionan a través de cálculos y procesos de computación. Su mecanismo consiste en imitar el lenguaje mediante la réplica de datos almacenados y el uso de patrones de probabilidad, lo que hace que su funcionamiento dependa totalmente de la calidad de los datos utilizados para su entrenamiento (Pereira Ginet-Jaquemet, 2024).

El análisis práctico realizado a partir del testimonio de Kyle Rittenhouse ha permitido contrastar esta cuestión desde una perspectiva empírica. Aunque DeepL Voice consiguió transmitir el sentido general del discurso y mantener cierta continuidad durante gran parte de la interacción, también mostró limitaciones importantes en aspectos especialmente sensibles dentro del ámbito judicial. Los errores terminológicos y contextuales observados, la simplificación de información potencialmente relevante desde un punto de vista probatorio, las dificultades para interpretar elementos pragmáticos del discurso o la incapacidad para gestionar adecuadamente situaciones de fuerte carga emocional muestran que la fidelidad del testimonio todavía no puede garantizarse plenamente mediante sistemas automatizados.

Estas limitaciones adquieren especial relevancia si se tiene en cuenta que, tal y como señala Lee (2015), el intérprete judicial no solo debe trasladar el significado proposicional del discurso, sino también preservar el registro, el estilo y el tono del hablante (Lee, 2015, p. 223). Asimismo, el contenido, la coherencia y la forma de expresión del testimonio pueden influir directamente en la percepción de credibilidad de una persona dentro del proceso judicial (Lee, 2015, p. 232). En este sentido, los resultados del análisis muestran que *DeepL Voice* encuentra mayores dificultades precisamente donde la dimensión humana del lenguaje resulta más determinante: las vacilaciones, el nerviosismo, el llanto, las interrupciones o los cambios de entonación. Lejos de interpretar estos elementos como parte del testimonio, la herramienta tiende a tratarlos de manera mecánica, generando secuencias poco naturales o incoherentes.

A ello se suma la ausencia de conciencia del error, un problema especialmente delicado desde el punto de vista procesal. Mientras que el intérprete humano posee la capacidad de detectar pérdidas de sentido, solicitar aclaraciones o incluso interrumpir la interpretación cuando la fidelidad del mensaje se ve comprometida —tal y como simbólicamente ya ocurría en Núremberg mediante el sistema de control de luces—, la inteligencia artificial continúa generando *output* (respuestas automáticas) incluso cuando el *input* (información de entrada recibida) deja de ser plenamente inteligible. Como explican Goldsmith y Díaz (2024), la Traducción Automática del Habla funciona mediante un proceso secuencial de reconocimiento del habla, traducción y síntesis de voz, donde un error inicial puede multiplicarse progresivamente en un fenómeno conocido como *efecto cascada*. A ello se añade la dependencia de modelos entrenados con grandes cantidades de datos potencialmente sesgados o imperfectos, circunstancia que dificulta alcanzar los niveles de precisión exigidos en contextos jurídicamente complejos (Pereira Ginet-Jaquemet, 2023).

A la luz de la evidencia analizada, la primera pregunta que ha guiado este trabajo —si sería ético utilizar sistemas de inteligencia artificial en casos judiciales sensibles— requiere una respuesta prudente. En el estado actual de desarrollo, la utilización autónoma de estas

herramientas en procedimientos penales especialmente sensibles no parece compatible con las exigencias éticas y lingüísticas derivadas del derecho a un juicio justo. Como recuerda Lee (2015, p. 227), la mera provisión de un servicio de interpretación no garantiza por sí sola la equidad procesal si este carece de la calidad necesaria para preservar la integridad del testimonio. Del mismo modo, tanto la AIIC (2020) como el marco regulatorio europeo coinciden en considerar la aplicación de la inteligencia artificial en contextos judiciales como un ámbito de alto riesgo que exige supervisión humana constante (AI Act Service Desk, 2024).

No obstante, responder negativamente a esta primera cuestión no implica negar el valor que la inteligencia artificial puede aportar a la profesión interpretativa. La segunda pregunta planteada en este trabajo —hasta qué punto la inteligencia artificial podría asumir funciones del intérprete humano manteniendo la exactitud lingüística y respetando los estándares éticos— requiere una respuesta más matizada. La evidencia científica disponible muestra que estas tecnologías pueden desempeñar un papel útil como herramientas de apoyo, como, por ejemplo, preparación terminológica, gestión de glosarios, organización documental, localización rápida de información o asistencia cognitiva durante la interpretación (Pleșca et al., 2024). En este sentido, el impacto real de la inteligencia artificial parece estar menos orientado a sustituir al profesional y más a transformar sus métodos de preparación y trabajo.

En definitiva, el recorrido desde Núremberg hasta la inteligencia artificial demuestra que, aunque los instrumentos tecnológicos han evolucionado enormemente, la función esencial de la interpretación judicial permanece intacta. La innovación técnica puede modificar la manera en que se presta el servicio, pero no elimina la necesidad de una mediación lingüística capaz de comprender el contexto, gestionar la emoción, interpretar la intención del hablante y asumir responsabilidad sobre las consecuencias del mensaje transmitido. La historia de Núremberg enseñó que, cuando el lenguaje puede condicionar el destino de una persona, la precisión interpretativa constituye una garantía de justicia. Ocho décadas después, esa lección sigue plenamente vigente. La inteligencia artificial puede consolidarse como una aliada valiosa de la profesión, pero, por el momento, no parece preparada para sustituir de manera autónoma la responsabilidad interpretativa que exige el ámbito judicial penal.

7.1. Limitaciones del estudio y perspectivas de futuro

Si bien este trabajo ha permitido contrastar la viabilidad de la inteligencia artificial en contextos judiciales, es necesario señalar ciertas limitaciones. El análisis práctico se ha centrado en una herramienta de uso general, la cual permite observar el funcionamiento básico de estos sistemas, pero no agota la diversidad de plataformas profesionales existentes. No obstante, dado que todas estas herramientas comparten una arquitectura tecnológica similar, los fallos detectados, especialmente la falta de "conciencia del error" y la incapacidad para gestionar el tono

emocional, son estructurales y se repetirían en cualquier sistema de traducción automática actual (Goldsmith y Díaz, 2024).

Asimismo, es necesario precisar que la elección de *DeepL Voice* como herramienta de análisis fue el resultado de un proceso de investigación previo. Inicialmente, se intentó acceder a plataformas profesionales de interpretación remota, como Kudo, con el fin de obtener una perspectiva técnica más avanzada. No obstante, el acceso a demostraciones en dichas plataformas está restringido exclusivamente a entornos corporativos, y a pesar de haber realizado las gestiones a través de una empresa colaboradora, no se recibió respuesta en un plazo razonable. Esta barrera de acceso obligó a centrar el análisis empírico en soluciones de IA de consumo, lo cual, lejos de invalidar el estudio, refuerza la necesidad de una mayor transparencia en el sector de la tecnología aplicada a la justicia.

De cara a futuras investigaciones, es necesario seguir observando cómo evolucionan los modelos de lenguaje (LLM) y su integración en la interpretación. La cuestión principal ya no consiste únicamente en determinar si la inteligencia artificial puede sustituir al intérprete —algo que este trabajo considera ética y lingüísticamente inviable, dada la incapacidad de los sistemas actuales para gestionar la complejidad pragmática del discurso—, sino cómo los profesionales aprenderán a trabajar junto a estas máquinas. La investigación futura debería centrarse en definir cómo formar a los intérpretes para que actúen como supervisores críticos de la tecnología. El objetivo final es establecer protocolos donde la inteligencia artificial siempre esté bajo el control humano (el concepto de *human-in-the-loop*), garantizando que actúe como un apoyo y nunca como un sustituto que ponga en riesgo la justicia.

REFERENCIAS

AI Act Service Desk. (2024). Recital 61: High-risk AI systems in the administration of justice. European Commission. Recuperado el 12 de enero de 2026 de https://ai-act-service-desk.ec.europa.eu/en/ai-act/recital-61

AIIC. (2020, 4 de diciembre). It's a Kind of Magic? Artificial Intelligence and Human Insight [Webinar]. AIIC United Kingdom & Ireland Region. https://aiic.org/site/uk-ie/AI-interpreting

AIIC. (2021a, 8 de enero). How Soon Is Now? Automated Speech Translation Solutions and Devices [Webinar]. AIIC United Kingdom & Ireland Region. https://aiic.org/site/uk-ie/AI-interpreting

AIIC. (2021b, 15 de enero). Under Pressure. AI and Interpreting in the Future: Disruption, Perception and Reality [Webinar]. AIIC United Kingdom & Ireland Region. https://aiic.org/site/uk-ie/AI-interpreting

Baigorri-Jalón, J. (2000). La interpretación de conferencias: el nacimiento de una profesión. De Núremberg a las Naciones Unidas. Editorial Comares.

Barra, M. (2023). SeamlessM4T: Massively multilingual & multimodal machine translation. Meta AI Research. Recuperado el 5 de febrero de 2026 de https://ai.meta.com/research/publications/seamlessm4t-massively-multilingual-multimodal-machine-translation/

BoothMate. (s.f.). Frequently Asked Questions. Recuperado el 14 de febrero de 2026 de https://boothmate.app/home#boothmate-faq

Court TV. (2021, 10 de noviembre). *Rittenhouse Trial: Kyle Rittenhouse takes the stand* [Vídeo]. YouTube. https://www.youtube.com/watch?v=1nvaVlsUpzE

Gaiba, F. (1998). The Origins of Simultaneous Interpretation: The Nuremberg Trial. University of Ottawa Press.

Goethe-Institut. (2021). Los intérpretes de Núremberg [Exposición y ciclo de conferencias]. Material documental y notas de presentación. Madrid, España.

Goldsmith, J. (2024a, 17 de enero). Make Your Interpreting Glossaries Pop With Visuals, Audio, and Links. Techforword. https://www.techforword.com/blog/multimedia-glossaries

Goldsmith, J. (2024b, 19 de febrero). AI-Powered Voices: A Game-Changer for Productivity and Language Learning. Techforword. https://www.techforword.com/blog/ai-powered-voices

Goldsmith, J. (2025a). Guide To Help You Talk to Interpreting Clients About AI. Techforword. https://www.techforword.com/blog/ai-decision-tree

Goldsmith, J. (2025b, 11 de diciembre). NotebookLM: The Language Professional's Secret Weapon for Conquering Information Overload. Techforword. https://www.techforword.com/blog/notebooklm-for-interpreters

Goldsmith, J. y Díaz, N. (2024). Machine Interpreting Showdown [Seminario web]. Techforword. https://www.techforword.com/blog/machine-interpreting-showdown

International Military Tribunal. (2021). The Nuremberg Trials: Volume 1 (Edición ePub). Musaicum Books. (Obra original publicada en 1947).

Lee, H. (2015). Court interpreting. In H. Mikkelson & R. Jourdenais (Eds.), The Routledge Handbook of Interpreting (pp. 222–237). Routledge.

Lu, J. y Fantinuoli, C. (2024). Computer-assisted interpreting tools and automation: A study on interpreter's cognitive load and automation complacency. Translation and Interpreting Studies, 19(2), 245–268. https://doi.org/10.1075/tis.23014.lu

Lu, X. y Fantinuoli, C. (2024). AI-assisted simultaneous interpreting: A study on the effect of terminology suggestions on performance and cognitive load. Linguistica Antverpiensia, New Series – Themes in Translation Studies, 22, 124–145. Recuperado el 20 de enero de 2026 de https://artojs02.uantwerpen.be/index.php/LANS-TTS/article/view/869

Mikkelson, H. (2015). Interpreting in the court. En H. Mikkelson y R. Jourdenais (Eds.), The Routledge handbook of interpreting (pp. 445–458). Routledge.

Pereira Ginet-Jaquemet, T. (2023). Exploring the challenges and realities of applying artificial intelligence to language interpretation. International Association of Conference Interpreters (AIIC). https://aiic.org/company/roster/companyRosterDetails.html?companyId=13594&companyRosterId=40

Plești, A.-M., Kusé, M., Woumans, E. y Defrancq, B. (2024, 2 de octubre). A comprehensive socio-cognitive study of simultaneous interpreting in an international institution: Report with focus on the Artificial Boothmate. Faculty of Arts and Philosophy, Ghent University / European Commission DG SCIC. https://backoffice.biblio.ugent.be/download/01JSH7XNPH4P2CXNW6573ATYAA/01JSH821P2AZGQCRJHK5A0QGKN

Pöchhacker, F. (2016). Introducing Interpreting Studies (2.ª ed.). Routledge.

Roland, R. A. (1999). Interpreters as diplomats: A diplomatic history of the role of interpreters in world politics. University of Ottawa Press.

Stansfield, C. W. (Ed.). (1990). Interpreting—Yesterday, today, and tomorrow. American Translators Association Scholarly Monograph Series. State University of New York Press.

Wordly. (s.f.). Interpretación simultánea de IA para reuniones y eventos. Recuperado el 22 de febrero de 2026 de https://www.wordly.ai/es/simultaneous-interpretation

ANEXOS

ANEXO I: TRANSCRIPCIONES

1. Transcripción original (inglés)

Lawyer: “Describe your approach to Car Source number three.”

Kyle Rittenhouse: “As I'm walking down Sheridan Road, I hear somebody scream, ‘Burn in hell.’”

Kyle Rittenhouse: “And I reply with, ‘Friendly, friendly, friendly.’ To let them know, hey, I'm just here to help.”

Kyle Rittenhouse: “I don't want any problems. I just want to put out the fires if there are any.”

Kyle Rittenhouse: “I continue walking and then I notice the Duramax, I notice a flame in the back seat of the Duramax…”

Kyle Rittenhouse: “As I'm stepping forward, I believe his name is now Joshua Ziminski, he steps towards me with a pistol in his hand.”

Kyle Rittenhouse: “As I'm walking towards to put out the fire, I drop the fire extinguisher and I take a step back.”

Lawyer: “When you step back from Mr. Ziminski, what's your plan?”

Kyle Rittenhouse: “My plan is to get out of that situation and go back north down Sheridan Road to where the Car Source lot number two was.”

Lawyer: “And did you get back, were you able to go in a northerly direction?”

Kyle Rittenhouse: “I wasn't.”

Lawyer: “Describe what happens.”

Kyle Rittenhouse: “Once I take that step back, I look over my shoulder and Mr. Rosenbaum was now running from my right side…”

Kyle Rittenhouse: “I was cornered from in front of me with Mr. Ziminski and there were three people right there.”

Lawyer: “Take a deep breath, Kyle.”

Judge: “We’re going to just take a… Time for our break anyway. You can just relax for a minute, sir.”

Judge: “We’re going to take a break about 10 minutes and please don't talk about the case during the break, read, watch and listen to any comment on the trial.”

2. Transcripción generada por DeepL Voice

Abogado: “Describe tu enfoque respecto a la fuente de datos número tres.”

Kyle Rittenhouse: “Mientras camino por Sheridan Road, oigo a alguien gritar: ‘¡Arde en el infierno!’”

Kyle Rittenhouse: “Y yo respondí con amabilidad, amabilidad, amabilidad para hacerles saber: ‘Oye, solo estoy aquí para ayudar.’”
Kyle Rittenhouse: “Solo es que no quiero problemas. Solo quiero apagar los incendios, si es que hay alguno.”
Kyle Rittenhouse: “Sigo caminando y entonces me doy cuenta del Duramax. Me fijé en una llama en el asiento trasero del Duramax…”
Kyle Rittenhouse: “Y mientras avanza, creo que ahora se llama Joshua Zaminski. Él se acercó a mí con una pistola en la mano.”
Kyle Rittenhouse: “Mientras caminaba, mientras caminaba hacia allí para apagar el fuego, se me cayó el extintor y di un paso atrás.”
Abogado: “Cuando das un paso atrás del Sr. Zeminski, ¿cuál es su plan?”
Kyle Rittenhouse: “Mi plan es salir de esa situación y volver al norte, a Sheridan Road, donde está el aparcamiento de coches número dos.”
Abogado: “¿Y consiguió volver? ¿Pudiste dirigirte hacia el norte?”
Kyle Rittenhouse: “No lo hice.”
Abogado: “Describa lo que ocurrió.”
Kyle Rittenhouse: “Una vez que di ese paso atrás. Miro por encima del hombro y… Sr. Rosenbaum, Sr. Rosenbaum ahora corría por mi lado derecho.”
Kyle Rittenhouse: “Y yo estaba acorralado por delante con el Sr. Zeminski. Y había… había gente justo ahí. Eso es lo que yo... Eso es lo que traje.”
Juez: “Vamos a tomarnos un momento para el descanso, de todos modos, puedes relajarte un momento, señor.”
Juez: “Vamos a hacer una pausa de unos 10 minutos, y por favor, no hable del caso durante la pausa. Lea, vea o escuche cualquier comentario.”

3. Propuesta de interpretación humana de referencia (español)
Abogado: “Describa qué ocurrió cuando se acercó al aparcamiento de Car Source número tres.”
Kyle Rittenhouse: “Mientras caminaba por Sheridan Road, oí a alguien gritar: ‘¡Arde en el infierno!’”
Kyle Rittenhouse: “Y yo respondí: ‘Tranquilos, tranquilos, tranquilos’, para hacerles saber: ‘Oye, solo estoy aquí para ayudar.’”
Kyle Rittenhouse: “No quería problemas. Solo quería apagar cualquier incendio, si lo había.”
Kyle Rittenhouse: “Seguí caminando y vi el Duramax; vi una llama en el asiento trasero…”
Kyle Rittenhouse: “Mientras avanzaba, creo que su nombre es Joshua Ziminski, se acercó hacia mí con una pistola en la mano.”
Kyle Rittenhouse: “Mientras me acercaba para apagar el fuego, solté el extintor y di un paso atrás.”
Abogado: “Cuando se alejó del Sr. Ziminski, ¿cuál era su intención?”

Kyle Rittenhouse: "Mi intención era salir de esa situación y regresar hacia el norte por Sheridan Road, donde se encontraba el aparcamiento de Car Source número dos."
Abogado: "¿Pudo regresar? ¿Consiguió dirigirse hacia el norte?"
Kyle Rittenhouse: "No pude."
Abogado: "Describa qué ocurrió después."
Kyle Rittenhouse: "Después de dar ese paso atrás, miré por encima del hombro y el Sr. Rosenbaum venía corriendo desde mi lado derecho."
Kyle Rittenhouse: "Estaba acorralado, con el Sr. Ziminski delante de mí, y había tres personas justo allí. *El acusado llora y no puede continuar.*"
Abogado: "Respire hondo, Kyle."
Juez: "Vamos a hacer una pausa. De todos modos, ya es hora del descanso. Puede relajarse un momento, señor."
Juez: "Vamos a hacer una pausa de unos diez minutos y, por favor, no hable del caso durante la pausa ni lea, vea o escuche comentarios sobre el juicio."

ANEXO II: CÓDIGO DE ACCESO Y REFERENCIA MULTIMEDIA

El análisis práctico realizado en este trabajo se fundamenta en el testimonio de Kyle Rittenhouse durante su juicio penal. Para la evaluación de la herramienta *DeepL Voice*, se ha seleccionado el fragmento comprendido entre el **minuto 1:33 y el 4:14** del vídeo original publicado por *Court TV* (2021).

Enlace al vídeo: https://www.youtube.com/watch?v=1nvaVlsUpzE

A continuación, se facilita el acceso al vídeo con la **interpretación automatizada** generada para este estudio, la cual sirve como prueba empírica de las limitaciones detectadas:

Enlace directo al vídeo: https://drive.google.com/file/d/1EJ0OxF1Mkp7cpkwrsBIzaK31M-Vv6RTz/view?usp=sharing

Acceso mediante código QR: