



UNIVERSIDAD DE MÁLAGA



GRADO EN INGENIERÍA DE LA SALUD

DISEÑO E IMPLEMENTACIÓN DE UN MÓDULO PARA EL ESTUDIO DE
TAREAS VISUALES Y RESPUESTAS ORALES A ESTÍMULOS DINÁMICOS.
OBTENCIÓN DE INDICADORES DE DESEMPEÑO DIGITALES Y SU
RELACIÓN CON LA EDAD.

DESIGN AND IMPLEMENTATION OF A MODULE FOR STUDYING VISUAL
TASKS AND ORAL RESPONSES TO DYNAMIC STIMULI. EXTRACTION OF
DIGITAL PERFORMANCE METRICS AND THEIR RELATIONSHIP WITH AGE.

Realizado por

AIMAR BARRANCO LOUKILI

Tutorizado por

GONZALO JOYA CAPARRÓS

Departamento

MÁLAGA, a FEBRERO de 2026



UNIVERSIDAD
DE MÁLAGA



ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA INFORMÁTICA
GRADUADO EN INGENIERÍA DE LA SALUD

DISEÑO E IMPLEMENTACIÓN DE UN MÓDULO PARA EL ESTUDIO DE
TAREAS VISUALES Y RESPUESTAS ORALES A ESTÍMULOS DINÁMICOS.
OBTENCIÓN DE INDICADORES DE DESEMPEÑO DIGITALES Y SU RELACIÓN
CON LA EDAD.

DESIGN AND IMPLEMENTATION OF A MODULE FOR STUDYING VISUAL
TASKS AND ORAL RESPONSES TO DYNAMIC STIMULI: EXTRACTION OF
DIGITAL PERFORMANCE METRICS AND THEIR RELATIONSHIP WITH AGE.

Realizado por

AIMAR BARRANCO LOUKILI

Tutorizado por

GONZALO JOYA CAPARRÓS

Departamento

TECNOLOGÍA ELECTRÓNICA



UNIVERSIDAD
DE MÁLAGA



UNIVERSIDAD DE MÁLAGA
MÁLAGA, FEBRERO DE 2026

Fecha defensa: (Marzo)

Resumen

Este Trabajo de Fin de Grado presenta el diseño y validación de un protocolo experimental software para cuantificar, de forma no invasiva, el desempeño del usuario en una tarea de seguimiento visomotor y en una tarea de respuesta verbal ante estímulos controlados. A diferencia de aproximaciones basadas en el seguimiento visual, en este estudio no se registran movimientos oculares: la variable principal de seguimiento se obtiene a partir de trayectoria del cursor mientras el participante intenta mantenerlo sobre un estímulo visual en movimiento, lo que permite evaluar de manera práctica la coordinación visomotora en un entorno reproducibles y de bajo coste.

A partir de los registros generados por la aplicación desarrollada, se calculan y analizan Indicadores de Desempeño (IDES) centrados en (i), el error espacial cursor-estímulo y su variabilidad, (ii) medidas de contacto/tiempo efectivo en objetivo, y (iii) el desempeño verbal asociado a cambios de color (latencias y aciertos). El análisis estadístico se plantea sobre una muestra de población general, poniendo el foco en la edad como variable explicativa, con el objetivo de identificar posibles tendencias compatibles con cambios funcionales esperables en la coordinación sensoriomotora y en la componente cognitivo-motora, sin establecer en ningún caso diagnósticos clínicos, y demostrar así, la validez del sistema para detección de posibles tendencias en la población.

Finalmente, se discuten limitaciones y líneas de mejora orientadas a ampliar la validez externa del protocolo y a su aplicación futura en estudios comparativos con líneas de investigación clínicas.

Palabras Clave: visomotor, seguimiento manual, cursor-estímulo, doble tarea, latencia verbal, envejecimiento, neurodegeneración, indicadores de desempeño, IDES.

Abstract

This Bachelor's Thesis presents the design and validation of a software-based experimental protocol to quantify user performance in a visuomotor tracking task and a speech response task under controlled stimuli. Unlike eye-tracking approaches, this work does not record eye movements: tracking performance is computed from the mouse cursor trajectory while the participant attempts to keep the cursor on a moving visual target, providing a practical, low-cost proxy to assess visuomotor coordination in a reproducible setting.

From the data collected by the developed application, a set of performance indicators (IDES) is computed, focusing on (i) spatial cursor–target error and its variability, (ii) on-target contact/effective time measures, and (iii) speech-related performance during colour changes (response latency and accuracy). Statistical analyses are conducted on a general-population sample, using age as the main explanatory variable to explore trends consistent with functional changes in sensorimotor coordination and cognitive–motor processing, without performing any clinical diagnosis.

Finally, limitations and future improvements are discussed to increase external validity and support future studies including clinical cohorts.

Keywords: visuomotor, manual tracking, cursor–target error, dual-task, speech latency, aging, neurodegeneration, performance indicators.

Índice

Contents

| | |
|--|-----------|
| Resumen..... | 1 |
| Abstract..... | 1 |
| Índice..... | 1 |
| Índice de Figuras..... | 5 |
| Índice de Tablas..... | 1 |
| Introducción..... | 3 |
| Motivación..... | 4 |
| Objetivos..... | 7 |
| Contribución a los objetivos de desarrollo sostenible (<i>Agenda 2030</i>)..... | 7 |
| Estructura de la memoria..... | 8 |
| Contexto..... | 11 |
| Estado del Arte..... | 11 |
| 2.1. Coordinación visomotora y seguimiento manual en envejecimiento y neurodegeneración..... | 11 |
| 2.2. Respuesta oral y análisis del habla en patologías cerebrales..... | 12 |
| Conocimiento biológico..... | 13 |
| 2.3. Visión-acción y coordinación visomotora..... | 13 |
| 2.4. Control motor de la mano y vías descendentes..... | 13 |
| 2.4.1. De la corteza motora a la mano por la vía corticoespinal..... | 13 |
| 2.4.2. Nervios periféricos y ejecución del movimiento..... | 14 |
| 2.4.3. Retroalimentación somatosensorial y corrección online..... | 14 |
| 2.5. Cerebelo, ganglios basales y estabilidad del control motor..... | 15 |
| 2.5.1. Cerebelo: predicción, ajuste fino y adaptación visomotora..... | 15 |
| 2.5.2. Ganglios basales: selección/iniciación y control del esfuerzo..... | 15 |
| 2.6. Producción del habla y su interacción con el control visomotor..... | 16 |
| 2.6.1. Bases neurofisiológicas de la respuesta verbal..... | 16 |
| 2.6.2. Doble tarea: por qué hablar puede degradar el seguimiento manual..... | 16 |
| 2.7. Relación con el envejecimiento y con el contexto neurodegenerativo..... | 17 |
| Tecnologías utilizadas..... | 18 |
| 2.8. Entorno de desarrollo y gestión de dependencias..... | 18 |
| 2.9. Interfaz de usuario y arquitectura modular..... | 18 |

| | |
|--|-----------|
| 2.10. Captura de audio y reconocimiento de voz con <i>Whisper</i> | 19 |
| 2.11. Sincronización y generación del <i>dataset</i> en formato <i>Parquet</i> | 19 |
| 2.12. Herramientas de análisis..... | 19 |
| Metodología y protocolo experimental..... | 20 |
| 2.13. Introducción..... | 20 |
| 2.14. Diseño general del estudio..... | 20 |
| 2.15. Participantes y consideraciones éticas..... | 20 |
| 2.16. Entorno experimental y entorno de ejecución..... | 21 |
| 2.17. Descripción de la tarea experimental..... | 22 |
| 2.18. Registro de señales visomotoras..... | 22 |
| 2.19. Captura y procesamiento de la respuesta verbal..... | 22 |
| 2.20. Sincronización multimodal y generación del conjunto de datos..... | 23 |
| 2.21. Preparación de los datos para el análisis..... | 24 |
| 2.22. Limitaciones metodológicas..... | 25 |
| Extracción de los IDES..... | 27 |
| 3.1. Introducción a los IDES..... | 27 |
| 3.2. Clasificación de los IDES..... | 28 |
| 3.3. IDES..... | 28 |
| 3.3.1. Error de seguimiento visomotor..... | 28 |
| 3.3.2. Variabilidad angular como IDE de estabilidad visomotora..... | 29 |
| 3.3.3. Número de contactos del cursor sobre el estímulo..... | 30 |
| 3.3.4. Tiempo efectivo en objetivo..... | 31 |
| 3.3.5. Latencia verbal ante el cambio de color del estímulo..... | 31 |
| 3.3.6. Precisión verbal..... | 32 |
| Resultados Obtenidos..... | 33 |
| 4.1. Descripción del conjunto de datos..... | 33 |
| 4.2. Flujo de procesamiento y extracción de IDES (resumen operativo)..... | 34 |
| 4.3. Comparabilidad entre ensayos..... | 35 |
| 4.4. Visión global de los usuarios..... | 38 |
| 4.5. IDE 1: Error de Seguimiento Visomotor..... | 39 |
| 4.6. IDE 2: Variabilidad angular (estabilidad direccional del seguimiento)..... | 42 |
| 4.7. IDE 3: Contactos del cursor sobre el estímulo (alcances exitosos)..... | 48 |
| 4.8. IDE 4: Tiempo efectivo en objetivo..... | 50 |
| 4.9. IDE 5: Latencia en la respuesta verbal al cambio de color..... | 53 |
| 4.10. IDE 6: Precisión verbal (aciertos/errores en denominación)..... | 59 |
| 4.11. Cierre de los Resultados Obtenidos..... | 64 |
| Conclusiones y líneas futuras..... | 67 |
| 5.1. Tendencias observadas..... | 68 |
| 5.2. Limitaciones del estudio..... | 68 |
| 5.3. Mejoras e implementaciones futuras..... | 69 |
| REFERENCIAS..... | 71 |
| Apéndice A. Manual de Instalación y Uso de <i>CHROMATRACKER</i>..... | 75 |
| A.1. Descripción general del sistema..... | 75 |
| A.2. Estructura del repositorio..... | 75 |
| A.3. Requisitos del sistema..... | 76 |
| i) Requisitos mínimos..... | 76 |

| | |
|---|----|
| ii) Requisitos recomendados..... | 76 |
| A.4. Instalación..... | 76 |
| i) Opción 1: Instalación con <i>uv</i> | 76 |
| ii) Opción 2: Instalación estándar con <i>venv + pip</i> | 76 |
| A.5. Ejecución del experimento..... | 77 |
| A.6. Archivos de salida y formato de datos..... | 77 |
| A.7. Cálculo de IDES con <i>Jupyter Notebook</i> y exportación a <i>Excel</i> | 78 |
| i) Apertura del entorno <i>Jupyter</i> | 78 |
| ii) Flujo general dentro del notebook..... | 78 |
| A.8. Solución de problemas frecuentes..... | 78 |
| i) Error: "No se encuentra la carpeta <i>parquetFinales</i> "..... | 78 |
| ii) La respuesta verbal no se registra o aparece vacía..... | 79 |
| iii) Errores con <i>sounddevice / PortAudio</i> | 79 |
| iv) Problemas con pantalla completa..... | 79 |
| A.9. Procedimiento recomendado para reproducir el flujo del TFG..... | 79 |

Índice de Figuras

| | |
|--|----|
| Figura 1: Integración visioacción durante el seguimiento manual. (Fuente: elaboración propia mediante IA generativa)..... | 13 |
| Figura 2: Vía motora descendente y bucles de corrección durante el control del cursor(Fuente: elaboración propia mediante IA generativa)..... | 15 |
| Figura 3: Modelo funcional de doble tarea: seguimiento manual + respuesta verbal (Fuente: elaboración propia mediante IA generativa)..... | 17 |
| Figura 4: Diagrama general de la arquitectura llevada a cabo(Fuente: elaboración propia mediante IA generativa)..... | 25 |
| Figura 5: Etapas del proceso de extracción y análisis de datos (Fuente: elaboración propia mediante IA generativa)..... | 34 |
| Figura 6: Error de seguimiento visomotor en función del tiempo (usuario 1)..... | 37 |
| Figura 7: Trayectorias del estímulo y del cursor durante el ensayo (usuario 1)..... | 37 |
| Figura 8: Tabla comparativa en z-score por usuario, ordenada por índice de atipicidad (media de z). Cada celda muestra la desviación relativa de una métrica respecto a la media del conjunto..... | 39 |
| Figura 9: Error de seguimiento (p95) por grupo de edad..... | 41 |
| Figura 10: Métrica angular (p95 abs) por grupo de edad..... | 44 |
| Figura 11: Serie temporal del ángulo en función del tiempo (ejecución con seguimiento estable).... | 45 |
| Figura 12: Serie temporal del ángulo en función del tiempo (ejecución con seguimiento inestable)46 | |
| Figura 13: Histograma del ángulo durante una ejecución con seguimiento estable..... | 46 |
| Figura 14: Histograma del ángulo durante una ejecución con seguimiento inestable..... | 47 |
| Figura 15: Relación entre el percentil 95 del ángulo absoluto y el percentil 95 del error de seguimiento..... | 48 |
| Figura 16: Número de contactos registrados por usuario..... | 50 |
| Figura 17: Fracción del tiempo en objetivo de cada usuario ordenados por edad..... | 52 |
| Figura 18: Media del tiempo en objetivo de cada..... | 52 |
| Figura 19: Latencias mostradas por grupo de edad ascendente en su percentil 95..... | 56 |
| Figura 20: Distribución de la latencia en su percentil 95 para el color rojo..... | 57 |
| Figura 21: Distribución de la latencia en su percentil 95 para el color amarillo..... | 57 |
| Figura 22: Distribución de la latencia en su percentil 95 para el color verde..... | 58 |
| Figura 23: Distribución media y mediana de la latencia por edad..... | 59 |
| Figura 24: Comparación directa de la precisión por registro de usuario (valores entre 0 y 1)..... | 61 |
| Figura 25: Precisión verbal por edad..... | 62 |
| Figura 26: Tasa de éxito para el color rojo..... | 62 |
| Figura 27: Tasa de éxito para el color amarillo..... | 63 |
| Figura 28: Tasa de éxito para el color verde..... | 63 |

Índice de Tablas

| | |
|---|----|
| Tabla 1: Lista de ensayos experimentales registrados..... | 34 |
| Tabla 2: Criterios de aceptación de los ensayos..... | 35 |
| Tabla 3: Flujo de inclusión/exclusión de pruebas..... | 36 |
| Tabla 4: Resumen de métricas de seguimiento visomotor para los 20 ensayos..... | 40 |
| Tabla 5: Variables empleadas en el cálculo del indicador IDE 2..... | 43 |
| Tabla 6: Número de contactos registrados por usuario..... | 49 |
| Tabla 7: Métricas usadas para el análisis del tiempo efectivo en objetivo..... | 51 |
| Tabla 9: Resumen de la latencia de respuesta verbal ante cambios de color..... | 54 |
| Tabla 10: Latencias separadas por color..... | 55 |
| Tabla 11: Cálculo de medias , para cada medida, de las latencias de respuesta verbal..... | 55 |
| Tabla 12: Variables empleadas en el cálculo de la precisión verbal..... | 60 |

1

Introducción

Este Trabajo de Fin de Grado describe el diseño e implementación de un sistema experimental en Python orientado a la captura y estructuración de datos multimodales durante una tarea controlada de seguimiento visomotor con un estímulo por pantalla. El objetivo técnico del trabajo es disponer de una infraestructura reproducible que permita registrar de forma coherente la interacción del usuario y derivar métricas cuantitativas de desempeño, en línea con iniciativas de investigación y docencia en tecnologías orientadas al estudio de señales y comportamiento en contextos de salud [1].

La tarea se fundamenta en estudios sobre coordinación ojo-mano y seguimiento continuo, donde el desempeño puede cuantificarse a partir de medidas de precisión, variabilidad y control motor durante la persecución del estímulo [2] [3] [4]. Además, hay evidencia de que el desempeño en tareas visomotoras y de seguimiento puede verse condicionado por variables como la edad y por alteraciones neurológicas, justificando el interés por disponer de protocolos y métricas comparables entre sujetos [5] [6] [7] [8] [9].

Haciendo alusión al planteamiento multimodal, el sistema contempla también un canal de respuesta verbal asociado a eventos controlados durante el ensayo (cambios de color del estímulo), permitiendo analizar el desempeño verbal en paralelo al seguimiento manual. Esta integración es coherente con

trabajos de doble tarea, y con aquellos artículos que vinculan el lenguaje con el deterioro neurocognitivo [10] [11].

En conjunto, el trabajo ofrece una infraestructura reproducible para registrar y analizar el desempeño durante tareas de seguimiento continuo, en coherencia con el marco de la coordinación ojo-mano, y el estudio cuantitativo de la persecución de un estímulo [2] [3] [4].

Motivación

Este Trabajo de Fin de Grado se centra en el interés creciente por desarrollar herramientas digitales accesibles, portables y de bajo coste que permitan cuantificar, de forma objetiva y reproducible, alteraciones leves tanto en el rendimiento motor como cognitivo. En el contexto de enfermedades neurodegenerativas y del envejecimiento, una detección temprana y su posterior seguimiento longitudinal, son especialmente relevantes debido a que los cambios funcionales pueden aparecer de manera progresiva y, sobretudo al principio, resulta difícil de identificar únicamente mediante observación clínica o pruebas aisladas [1] [2] [3].

En este ámbito, se valora mucho el uso de IDEs: métricas derivadas del desempeño de una persona en tareas controladas en un entorno computacional. Este tipo de aproximaciones permite medir fenómenos como la coordinación ojo-mano, la variabilidad del movimiento, u otras capacidades cognitivas que pueden llegar a decaer por causas como el envejecimiento o alteraciones neurológicas [5] [6] [7] [8] [9] [10] [11].

Sin embargo, existe una limitación práctica debido a que muchas soluciones disponibles dependen de equipos especializados, configuraciones complejas o entornos poco reproducibles, lo que hace más difícil su implantación en contextos ambulatorios, descentralizados o con recursos limitados [1].

A partir de esta necesidad surge la motivación principal de este TFG: construir una infraestructura experimental en Python capaz de generar *datasets* multimodales, sincronizados y exportables, que permitan extraer de forma sistemática IDEs a partir de una tarea sencilla, repetible y controlada. El objetivo es capaz de proporcionar una plataforma que, sin requerir instrumentación clínica avanzada, sea capaz de registrar de manera coordinada: [1]

- La evolución respecto al tiempo del estímulo visual mostrado por pantalla,
- La trayectoria del cursor durante el seguimiento, y
- La respuesta verbal del usuario mediante captura y transcripción automática,

integrando además un mecanismo de sincronización temporal y una exportación estructurada que facilite el análisis posterior [1].

El problema que aborda este estudio no es sólo "medir" una variable concreta, sino resolver una dificultad metodológica recurrente: cómo obtener datos multimodales consistentes y comparables entre sujetos, con calidad suficiente como para estudiar relaciones entre métricas y analizar tendencias poblacionales. Para ello, el sistema desarrollado implementa una cadena completa que abarca desde el control del estímulo y el registro de eventos, hasta el cálculo de IDES robustos (incluyendo métricas basadas en percentiles para capturar episodios críticos), permitiendo reproducibilidad y trazabilidad del experimento [5] [7] [9] [10].

Dado que el acceso a poblaciones clínicas está limitado, este trabajo emplea datos correspondientes a sujetos sanos de diferentes edades, ayudando así a estudiar cómo el envejecimiento se asocia a cambios en el desempeño visomotor y verbal en una tarea controlada, y además, desarrollar una referencia metodológica para futuros estudios donde se comparen patrones de desempeño entre sujetos sanos y poblaciones con alteraciones neurológicas [7] [8] [9] [10].

De esta manera, la contribución del TFG se centra en la plataforma y procedimiento de extracción, así como en la definición y análisis de un conjunto de IDES que describen el rendimiento del usuario en varios ámbitos (precisión y estabilidad del seguimiento, dinámica de contacto con el objetivo, y desempeño en la respuesta verbal) [10][11].

En resumen, la motivación se centra en aportar una solución práctica y reproducible para capturar, sincronizar y analizar rendimiento multimodal en tareas de seguimiento y denominación, facilitando la generación de evidencia cuantitativa que pueda servir como base para futuras investigaciones o extensiones del sistema en contextos tanto experimentales como clínicos [1] [2] [3] [7].

Objetivos

El objetivo principal de este Trabajo Fin de Grado es diseñar, implementar y validar un sistema experimental multimodal que permita registrar y analizar el desempeño visomotor y cognitivo-verbal de usuarios sanos durante una tarea de seguimiento visomotor con respuesta verbal simultánea, mediante la definición y estudio de IDES cuantitativos.

Para alcanzar el objetivo general planteado, se establecen los siguientes objetivos específicos:

1. Diseñar una tarea experimental controlada que combine el seguimiento visomotor de un estímulo dinámico con una respuesta verbal ante eventos discretos, garantizando condiciones homogéneas de ejecución.
2. Desarrollar una aplicación software capaz de registrar de forma sincronizada la posición del estímulo, la posición del cursor y la respuesta verbal del usuario durante la ejecución de los ensayos.
3. Definir y justificar un conjunto de IDES que permitan cuantificar de forma objetiva distintos aspectos del desempeño visomotor y cognitivo-verbal durante la tarea experimental.
4. Implementar los procedimientos necesarios para el cálculo de los IDES a partir de los datos registrados, asegurando su reproducibilidad y coherencia con el diseño experimental.
5. Aplicar el sistema desarrollado en un estudio experimental con población sana, obteniendo un conjunto de datos que permita analizar el comportamiento de los indicadores bajo condiciones controladas.
6. Analizar de forma exploratoria los resultados obtenidos, evaluando la variabilidad de los IDES entre ensayos y usuarios, y discutiendo su utilidad como métricas funcionales en el contexto del experimento realizado.

Contribución a los objetivos de desarrollo sostenible (*Agenda 2030*)

Este Trabajo Fin de Grado contribuye de manera indirecta a varios de los *Objetivos de Desarrollo Sostenible (ODS)* establecidos por las Naciones Unidas, especialmente en aquellos relacionados con la salud, la educación y el desarrollo tecnológico accesible. [1]

En primer lugar, el trabajo se alinea con el **ODS 3**:

Salud y bienestar, al abordar el estudio de funciones visomotoras y cognitivo-verbales mediante una metodología experimental no invasiva. Aunque el estudio se ha realizado sobre población sana y con

un enfoque exploratorio, el desarrollo de herramientas que permitan cuantificar de forma objetiva el desempeño funcional puede resultar de interés en líneas de investigación orientadas al seguimiento y evaluación de procesos neurológicos. La definición de indicadores cuantitativos y reproducibles contribuye a la generación de conocimiento en el ámbito de la ingeniería aplicada a la salud.

Asimismo, el proyecto guarda relación con el **ODS 4**:

Educación de calidad, al integrar conocimientos de programación, procesamiento de señales, diseño experimental y análisis de datos dentro de un mismo marco de trabajo. El sistema desarrollado y la metodología propuesta pueden emplearse como recurso formativo en entornos académicos, facilitando la enseñanza de conceptos relacionados con la adquisición y el análisis de datos multimodales en el ámbito de la ingeniería biomédica y de la salud.

Por otro lado, el uso de tecnologías software de código abierto y de dispositivos de entrada convencionales contribuye al **ODS 9**:

Industria, innovación e infraestructura, al fomentar el desarrollo de soluciones tecnológicas accesibles, reproducibles y de bajo coste. Este enfoque favorece la replicabilidad del experimento y la reutilización del sistema en contextos de investigación o formación sin requerir equipamiento especializado.

Estructura de la memoria

La presente memoria se organiza de forma progresiva para conducir al lector desde la motivación del problema hasta la validación del sistema y el análisis de los indicadores obtenidos.

En primer lugar, el capítulo de **Introducción** integra la introducción general, la **Motivación**, los **Objetivos**, la **Contribución a los objetivos de desarrollo sostenible (Agenda 2030)** y el propio apartado de **Estructura de la memoria**, con el fin de contextualizar el propósito del trabajo, justificar la necesidad de una herramienta experimental reproducible y delimitar su alcance.

Seguidamente, el capítulo de **Contexto** establece el marco teórico y técnico del TFG, incluyendo el **Estado del Arte** en coordinación visomotora, seguimiento manual y habla en envejecimiento/neurodegeneración; las bases de **Conocimiento biológico** necesarias para interpretar la tarea y las métricas; y las **Tecnologías utilizadas** para implementar la aplicación, registrar señales multimodales, sincronizarlas y preparar el análisis. Dentro de este mismo capítulo se incorpora la sección de **Metodología y protocolo experimental**, donde se detalla el diseño del estudio, el entorno de ejecución, la tarea experimental, el registro visomotor y verbal, la sincronización multimodal, la generación del *dataset* y los criterios de preparación y validación de los ensayos.

A continuación, el capítulo de **Extracción de los IDES** define y formaliza los indicadores de desempeño empleados, describiendo su fundamento y el procedimiento de cálculo a partir de los registros generados por la aplicación.

El capítulo de **Resultados obtenidos** presenta el conjunto de datos analizado, el flujo de procesamiento y el estudio de cada IDE, apoyado en tablas y figuras, incluyendo comparaciones exploratorias en función de la edad y relaciones entre indicadores.

Finalmente, el capítulo de **Conclusiones y líneas futuras** sintetiza las aportaciones del trabajo, resume las tendencias observadas, discute limitaciones del protocolo y plantea mejoras orientadas a aumentar la validez externa y la aplicabilidad del sistema. La memoria se cierra con la sección de **Referencias** y un **Apéndice** que incluye el manual de instalación y uso de *ChromaTracker* para facilitar la reproducibilidad del flujo experimental y del análisis.

2

Contexto

Estado del Arte

2.1. Coordinación visomotora y seguimiento manual en envejecimiento y neurodegeneración

La coordinación visomotora es un proceso distribuido que integra percepción visual, atención, planificación motora y ejecución manual. En tareas de seguimiento continuo, el desempeño depende de la capacidad de transformar información visual en comandos motores, ajustar la trayectoria mediante retroalimentación y anticipar el movimiento del estímulo [5] [6]. La literatura sobre eye–hand coordination describe la participación coordinada de redes parieto–frontales y del cerebelo para sostener el acoplamiento espacial y temporal entre objetivo y respuesta manual, destacando que lesiones corticales suelen asociarse a errores espaciales, mientras que disfunciones cerebelosas afectan especialmente a la predicción temporal y la estabilidad del control motor [7].

En relación con el envejecimiento, se han reportado diferencias sistemáticas en tareas de interacción con ratón y control fino del cursor, observándose mayores dificultades en población de mayor edad en términos de precisión y eficiencia en tareas básicas de control [8].

Además, existe evidencia de que alteraciones visomotoras pueden aparecer de forma temprana en poblaciones con mayor riesgo de Alzheimer, especialmente cuando las tareas incrementan la demanda cognitivo–motora y la complejidad de la transformación visomotora [9].

En patologías como Parkinson también se han estudiado tareas de tracking visomotor con componente atencional, mostrando que el rendimiento puede degradarse cuando aumenta la distracción o la carga de control ejecutivo, lo que refuerza la utilidad de este tipo de paradigmas como pruebas funcionales sensibles a la interferencia cognitivo-motora [10] [3].

En conjunto, estos antecedentes justifican el interés por indicadores basados en el error cursor-estímulo, la estabilidad del seguimiento y la eficiencia de contacto con el objetivo como medidas prácticas para caracterizar desempeño visomotor en estudios por edad y como base para trabajos posteriores con cohortes clínicas.

2.2. Respuesta oral y análisis del habla en patologías cerebrales

El habla es otro dominio especialmente sensible a la afectación neurológica. La producción vocal depende de la coordinación entre la respiración, la fonación y la articulación, reguladas por redes que incluyen la corteza motora, áreas del lenguaje, el cerebelo, los ganglios basales y el tronco encefálico. Cuando estas estructuras se ven dañadas, aparecen alteraciones como la disartria, la variabilidad excesiva en el ritmo, la monotonía prosódica o el incremento en la latencia de respuesta [11] [14].

En el contexto de trastornos neurológicos con componente motor y/o cognitivo, tanto en trastornos predominantemente motores como en aquellos con un componente cognitivo relevante. Entre las alteraciones más habituales se encuentran la articulación imprecisa, los cambios en la prosodia, la aparición de pausas inadecuadas y la modificación anómala de la velocidad articulatoria [15].

El desarrollo de herramientas de análisis acústico ha permitido cuantificar de manera objetiva algunas de estas características [12]. Medidas como la frecuencia fundamental, la variabilidad del tono, la intensidad, la duración de las pausas o la estabilidad articulatoria se han utilizado para construir descriptores del estado del sistema motor del habla. No obstante, en contextos experimentales multimodales, también se han propuesto métricas más simples basadas en eventos, como la latencia de respuesta oral ante estímulos controlados (por ejemplo, tareas de denominación), que se interpreta como un posible Indicador de Desempeño relacionado con la planificación motora del habla y la velocidad de procesamiento cognitivo [14] [15] [16].

Conocimiento biológico

2.3. Visión–acción y coordinación visomotora

Como tarea central, este TFG se apoya en un principio de neurociencia motora: sistema nervioso transforma información sensorial (principalmente visual) en acciones motoras dirigidas a un objetivo. En una tarea de seguimiento continuo, el usuario debe (i) percibir la posición y velocidad del estímulo, (ii) compararlas con el estado actual del cursor, (iii) planificar una corrección motora y (iv) ejecutar esa corrección con la mano, repitiendo el ciclo de forma constante mientras el estímulo se mueve [17].

En términos funcionales, este proceso se explica mediante circuitos parietal–frontales responsables de las transformaciones sensoriomotoras: la información visual se integra con mapas espaciales y atencionales (corteza parietal posterior) y se conecta con áreas premotoras y motoras que generan planes de acción. Dicho proceso se ve reflejado en la Figura 1. Este acoplamiento visión–acción permite mantener la trayectoria de la mano alineada con el objetivo, anticipar su movimiento y corregir errores en tiempo real. En revisiones sobre circuitos parietal–frontales se describe precisamente esta transformación desde intención/planificación a ejecución, incluyendo estimación de estado y representación en coordenadas útiles para la acción [18].

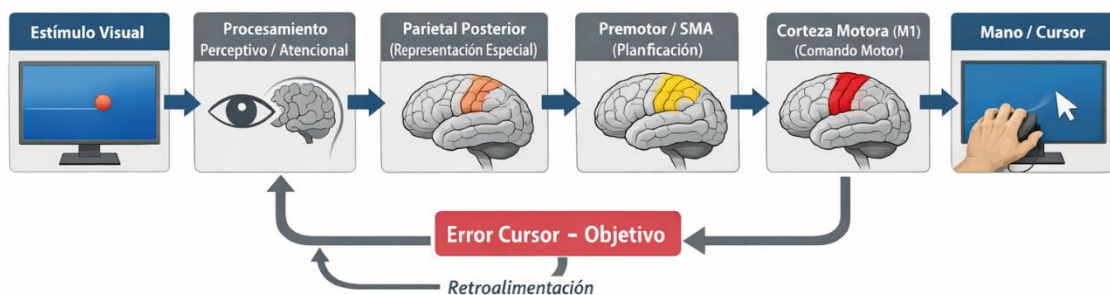


Figura 1: Integración visioacción durante el seguimiento manual. (Fuente: elaboración propia mediante IA generativa)

2.4. Control motor de la mano y vías descendentes

2.4.1. De la corteza motora a la mano por la vía corticoespinal

La ejecución del movimiento del cursor depende de la activación coordinada de músculos del hombro, brazo, antebrazo y mano. A nivel anatómico, el control voluntario fino de la mano está fuertemente

asociado a la vía corticoespinal, que conecta la corteza motora (M1 y regiones adyacentes) con circuitos medulares y motoneuronas que finalmente activan la musculatura distal. En humanos existe una considerable variabilidad individual en la organización de fibras motoras relacionadas con la mano dentro del tracto corticoespinal, lo que se ha estudiado con tractografía por difusión, reforzando la idea de que el control manual fino es un producto de redes descendentes con diferencias interindividuales medibles [19].

Desde el punto de vista funcional, la vía corticoespinal permite: iniciar y modular el movimiento voluntario, ajustar la fuerza y la dirección de la acción, y, especialmente, sostener control fino de la mano (microcorrecciones continuas), crucial en tareas de tracking [19].

2.4.2. Nervios periféricos y ejecución del movimiento

Aunque el comando motor se origina en el Sistema Nervioso Central, la ejecución final ocurre a través del sistema nervioso periférico. En la extremidad superior destacan tres nervios principales derivados del plexo braquial:

- Nervio mediano: importante en flexión de muñeca/dedos y pinza.
- Nervio cubital: control fino intrínseco de la mano.
- Nervio radial: extensión de muñeca/dedos y estabilización postural.

En una tarea con ratón, el movimiento se apoya tanto en musculatura proximal (estabilización de hombro y codo) como en control distal (muñeca/dedos), lo que hace que el desempeño sea sensible a variaciones en coordinación, fatiga y precisión.

2.4.3. Retroalimentación somatosensorial y corrección online

El seguimiento continuo no se sostiene solo con visión: la corrección fina depende de retroalimentación somatosensorial (propiocepción y tacto). Receptores como husos musculares (longitud/velocidad) y órganos tendinosos de Golgi (tensión), junto con la información cutánea de la mano, permiten ajustar la trayectoria y amortiguar temblores o sobrecorrecciones. En tracking continuo, estas entradas ayudan a minimizar oscilaciones y a estabilizar la trayectoria cuando el estímulo cambia de dirección o velocidad [2] [3] [18]. La Figura 2 muestra el desempeño de la vía motora descendente y sus correspondientes bucles de corrección durante el control del cursor.

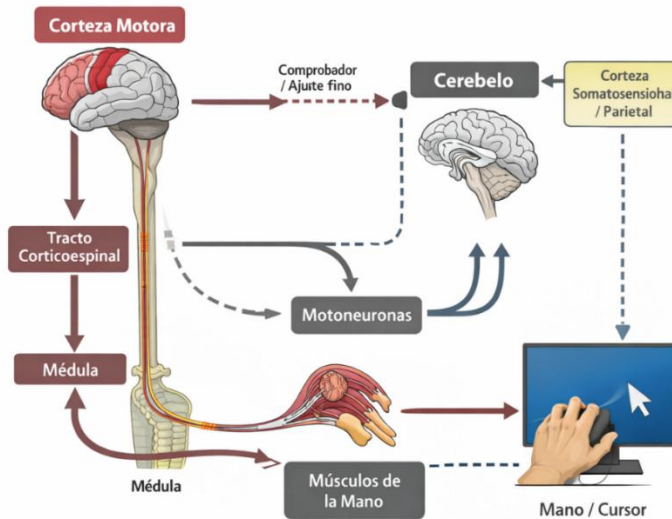


Figura 2: Vía motora descendente y bucles de corrección durante el control del cursor(Fuente: elaboración propia mediante IA generativa)

2.5. Cerebelo, ganglios basales y estabilidad del control motor

2.5.1. Cerebelo: predicción, ajuste fino y adaptación visomotora

En tareas de seguimiento, el cerebelo es especialmente relevante porque contribuye a: predicción del movimiento (anticipación), corrección de errores temporales, suavizado del movimiento, y ajuste continuo en presencia de cambios [20].

La literatura sobre adaptación visomotora destaca el papel central del cerebelo en el ajuste flexible de programas motores en respuesta a discrepancias entre lo esperado y lo observado, con interacción importante con corteza premotora y parietal [7].

2.5.2. Ganglios basales: selección/iniciación y control del esfuerzo.

Los ganglios basales participan en aspectos de selección/iniciación de acciones y control de la amplitud del movimiento, contribuyendo a la regulación del "esfuerzo motor" y de la consistencia de la respuesta. En tareas donde hay demanda sostenida y necesidad de mantener un patrón eficiente de seguimiento, alteraciones en estos circuitos suelen reflejarse en dificultad para iniciar ajustes, mantener estabilidad o sostener un rendimiento consistente bajo carga [7].

2.6. Producción del habla y su interacción con el control visomotor

2.6.1. Bases neurofisiológicas de la respuesta verbal

La respuesta verbal que se registra en el sistema no es solo "decir una palabra", implica una secuencia rápida de procesos:

1. percepción del cambio de estímulo (detección),
2. selección léxica (elección de la etiqueta verbal),
3. planificación motora del habla,
4. ejecución articulatoria (fonación y articulación).

A nivel cerebral, la producción del habla se sostiene en redes distribuidas (áreas frontales del lenguaje y corteza motora del habla), y está modulada por bucles con ganglios basales y cerebelo [16]. En trabajos que revisan la contribución de estas estructuras se describe que su daño o disfunción se asocia a perfiles específicos de alteración en habla/fluidez/planificación, lo que justifica el interés de métricas como latencia y precisión como indicadores funcionales [14].

A nivel periférico, la fonación depende del control de la laringe (cuerdas vocales) y de su coordinación con respiración y articulación [14].

2.6.2. Doble tarea: por qué hablar puede degradar el seguimiento manual

En el protocolo experimental se combina seguimiento continuo con la respuesta verbal. Desde el punto de vista fisiológico, esto introduce un componente de interferencia: ambos procesos compiten por recursos atencionales/ejecutivos y por la coordinación temporal de respuestas. Por eso, en un mismo usuario puede observarse el *trade-off* típico: mantener buena precisión de tracking, pero aumentar latencia verbal o responder rápido pero degradar estabilidad/precisión del cursor [21].

Esta interacción es precisamente la razón por la que en este trabajo se justifica un enfoque "multidimensional" (IDES visomotoras + verbales), ya que la degradación puede manifestarse en canales diferentes según el perfil del participante. En la Figura 3 se muestra dicho modelo de doble tarea.

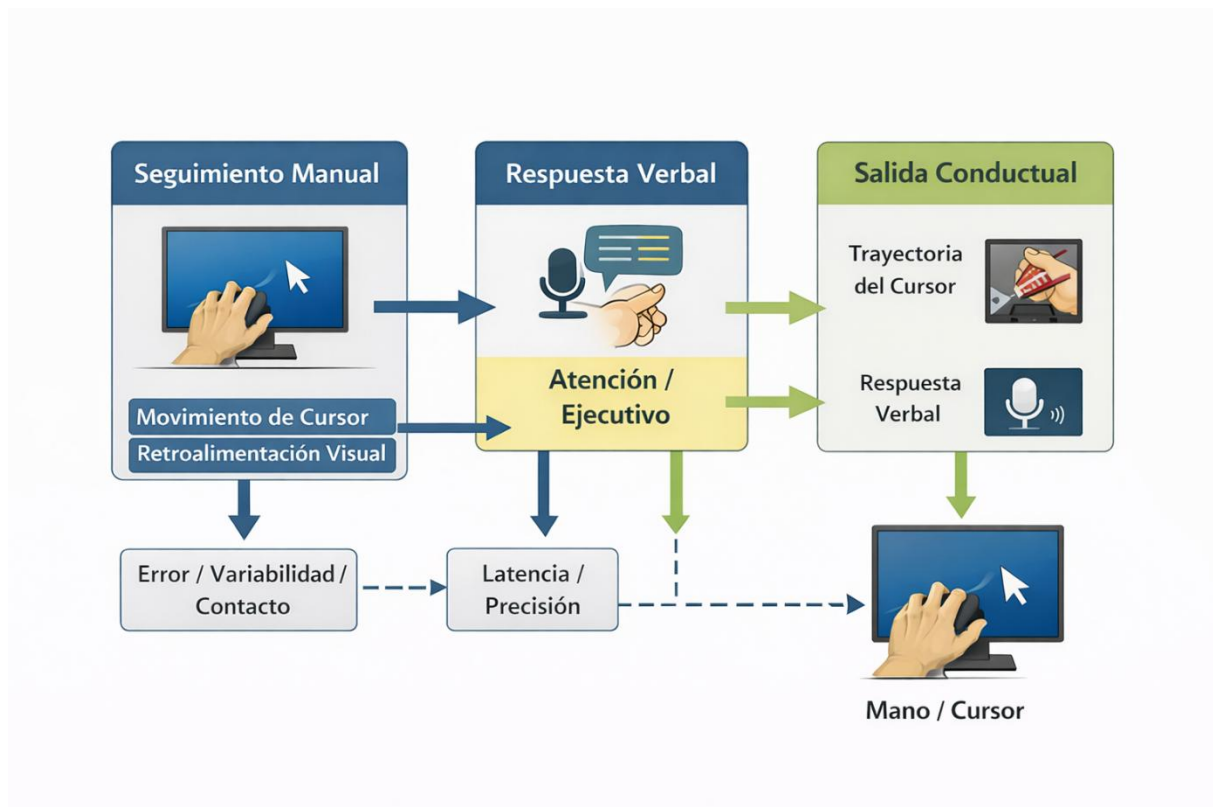


Figura 3: Modelo funcional de doble tarea: seguimiento manual + respuesta verbal (Fuente: elaboración propia mediante IA generativa)

2.7. Relación con el envejecimiento y con el contexto neurodegenerativo

El enfoque biológico anterior permite justificar por qué el análisis por edad es coherente con el objetivo del TFG. El envejecimiento afecta de forma gradual a la velocidad de procesamiento, la estabilidad del control motor fino, la capacidad de corrección online y la resistencia a la interferencia de doble tarea. Por ello, aun sin establecer diagnósticos, la caracterización cuantitativa de: precisión y estabilidad del tracking manual, eficiencia de contacto con el objetivo, y la latencia y precisión de denominación verbal, puede considerarse como un marco funcional plausible para estudiar tendencias asociadas a edad y para establecer una base de referencia que, en estudios futuros, facilite comparaciones con cohortes clínicas [5] [6] [7] [10] [8] [9].

Tecnologías utilizadas

2.8. Entorno de desarrollo y gestión de dependencias

La aplicación se desarrolló en [Python](#) dentro de un entorno aislado, con el objetivo de asegurar reproducibilidad y evitar conflictos entre versiones de librerías. Para la ejecución del software y el análisis posterior se utilizaron dependencias centradas en tres necesidades principales: (1) renderizado y control de la tarea experimental en tiempo real, (2) adquisición de audio para reconocimiento de voz, y (3) almacenamiento estructurado y sincronización de datos para el cálculo de indicadores.

Las dependencias principales empleadas en el sistema son: [Pygame](#) para la interfaz gráfica y el bucle de ejecución en tiempo real; [SoundDevice](#) para la captura de audio; [SciPy](#) para el guardado de audio en formato [WAV](#); [Pandas](#) y [Numpy](#) para la organización, sincronización y manipulación de señales; y [PyArrow](#) para la generación de ficheros en formato [Parquet](#). Para el análisis y representación en fase de estudio se emplearon [JupyterLab](#), [Matplotlib](#) y [OpenPyXL](#) (para exportación o lectura de resultados en Excel).

2.9. Interfaz de usuario y arquitectura modular

El proyecto se estructura en un fichero principal (*main.py*) y un paquete que agrupa la lógica del experimento y sus componentes. La arquitectura está diseñada de forma modular para separar responsabilidades y facilitar mantenimiento:

- Pantalla de instrucciones: implementada en *instructions.py*, muestra las indicaciones al usuario antes de iniciar el ensayo y refleja el estado de disponibilidad del modelo de reconocimiento de voz (por ejemplo, "loading" frente a "ready").
- Motor del experimento: *logic.py* coordina la secuencia completa de ejecución: inicialización, cuenta atrás, bucle principal de 60 segundos, finalización del ensayo y persistencia de datos.
- Estímulo visual: *ball.py* implementa la dinámica del estímulo (*posición, velocidad, rebotes, pausas y cambios de color*).
- Registro del cursor: *cursor.py* registra continuamente la posición del ratón como aproximación al seguimiento visual simulado.

Esta separación facilita describir el sistema como un entorno experimental reproducible, donde cada módulo aporta una parte concreta del registro multimodal.

2.10. Captura de audio y reconocimiento de voz con *Whisper*

La adquisición y procesado de voz se implementa mediante dos módulos:

- Carga del modelo en segundo plano (*whisper_loader.py*): la clase *WhisperLoader* lanza un hilo *daemon* que carga el modelo *OpenAI Whisper* sin bloquear la interfaz. En la versión validada del proyecto el modelo se carga como *load_model* ("medium", device="cpu"), garantizando ejecución en equipos sin GPU.
- Grabación y extracción de eventos verbales (*whisper.py*): la clase *SpeechRecorder* captura audio durante la duración completa del ensayo mediante *SoundDevice* y guarda temporalmente un WAV usando *SciPy*. Al finalizar, se transcribe el audio usando el modelo *Whisper* ya cargado, y se procesa la transcripción para identificar específicamente las palabras objetivo ("rojo", "amarillo", "verde"). Tras normalización del texto (minúsculas, eliminación de acentos y filtrado de caracteres), cada color se codifica numéricamente (1, 2, 3) para su posterior análisis y sincronización.

2.11. Sincronización y generación del *dataset* en formato *Parquet*

Para el almacenamiento final se emplea el formato *Parquet*, por su eficiencia y compatibilidad con flujos de análisis en *Python*. Los registros del estímulo, del cursor y los eventos de voz se integran en un *dataset* sincronizado mediante *Pandas*, utilizando una fusión temporal que preserva coherencia entre señales de distinta naturaleza.

El fichero final por ensayo contiene, de forma alineada temporalmente, variables como: *timestamp*, posición del estímulo (*ball_x*, *ball_y*), *ball_color_code*, posición del cursor (*cursor_x*, *cursor_y*) y el código verbal *voice_color_code*. Este archivo se guarda como *final_XXX.parquet*, eliminándose los ficheros temporales de trabajo.

2.12. Herramientas de análisis

El cálculo de indicadores y la obtención de tablas y gráficas se realizaron principalmente en *JupyterLab*, empleando *Pandas/NumPy* para el procesado numérico y *Matplotlib* para la visualización. La exportación e integración de resultados en formatos tabulares se realizó con soporte de *OpenPyXL*, permitiendo trabajar con hojas Excel en la fase de análisis final.

Metodología y protocolo experimental

2.13. Introducción

En este capítulo se describe la metodología seguida para la ejecución de las pruebas experimentales realizadas en este Trabajo Fin de Grado, así como el protocolo aplicado para la adquisición de los datos empleados posteriormente en el cálculo de los IDES. El objetivo principal de este apartado es detallar, de forma clara y reproducible, cómo se ha diseñado y ejecutado la tarea experimental, qué señales se han registrado y en qué condiciones se han obtenido los datos analizados en los capítulos posteriores.

La metodología propuesta se apoya en una tarea multimodal desarrollada específicamente para este proyecto, que combina seguimiento visomotor simulado de un estímulo dinámico y respuesta verbal ante cambios controlados del estímulo. Esta aproximación permite obtener, en un único experimento, información visomotora y cognitivo-verbal sincronizada, utilizando únicamente un ordenador convencional, un ratón y un micrófono, sin necesidad de equipamiento especializado.

2.14. Diseño general del estudio

El estudio se plantea como un estudio experimental exploratorio, aplicado inicialmente en población sana, con el propósito de validar el funcionamiento del sistema desarrollado y analizar el comportamiento de los IDES en función de la edad. Al no disponer de participantes diagnosticados con enfermedades neurodegenerativas, los resultados obtenidos deben interpretarse como valores de referencia funcionales, útiles para establecer comparaciones en futuros estudios con población clínica.

El diseño experimental se basa en la repetición de una tarea estándar de duración fija, ejecutada bajo condiciones controladas y homogéneas. Cada ejecución genera un conjunto de datos sincronizados que recoge la evolución temporal del estímulo visual, la respuesta visomotora del usuario y la respuesta verbal ante eventos discretos.

2.15. Participantes y consideraciones éticas

Las pruebas se realizaron con participantes sanos, sin diagnóstico conocido de patología neurológica o trastornos del habla. Todos los sujetos participaron de forma voluntaria y realizaron la tarea bajo las mismas condiciones experimentales. Se adjuntan a petición de aquel que quiera, unas hojas de

consentimiento donde cada usuario ha firmado el tratamiento de los datos obtenidos en su correspondiente prueba.

En cumplimiento de la normativa vigente en materia de protección de datos y privacidad, no se almacenó ningún dato personal identificativo de los participantes. Los registros generados se limitaron exclusivamente a información funcional derivada de la interacción con el sistema (trayectorias, marcas temporales y códigos de respuesta), garantizando en todo momento la anonimización de los datos. Asimismo, no se conservaron grabaciones de audio crudas tras el procesado, eliminándose los archivos temporales una vez finalizada la transcripción, de acuerdo con los principios de minimización de datos.

2.16. Entorno experimental y entorno de ejecución

Con el objetivo de minimizar la variabilidad introducida por el entorno y el hardware, todas las pruebas experimentales se realizaron bajo condiciones físicas homogéneas. Los ensayos se llevaron a cabo en un mismo ordenador, utilizando siempre el mismo dispositivo apuntador (*ratón inalámbrico de 2.4 GHz*) y manteniendo una configuración fija de pantalla.

El monitor se dispuso en orientación vertical, colocado sobre una mesa de altura constante (≈ 74 cm), y el participante permaneció sentado con una postura estable durante toda la ejecución del ensayo. La distancia ojo–pantalla se mantuvo aproximadamente constante en torno a 70 cm. Esta distancia se sitúa dentro del rango recomendado por guías ergonómicas generales, como las indicadas por la OSHA [22], que sugieren una distancia de visualización entre 50 y 100 cm, y por la Mayo Clinic, que recomienda situar el monitor aproximadamente a "*distancia de un brazo*" [23].

Asimismo, la configuración del sistema se mantuvo constante durante todas las pruebas, verificándose una resolución de 1920×1080 píxeles y una frecuencia de refresco fija de 60 Hz, sin variación dinámica de la tasa de actualización. Esta configuración es coherente con la tasa objetivo de ejecución de la aplicación (60 FPS) y evita cambios en la presentación del estímulo entre ensayos. Este conjunto de condiciones contribuye a reducir la variabilidad asociada a factores posturales, ergonómicos y de visualización, asegurando que las diferencias observadas en los IDES se deban principalmente al desempeño del usuario y no a cambios en el entorno experimental.

Esta configuración responde al objetivo del proyecto de desarrollar una herramienta experimental accesible y fácilmente reproducible en distintos contextos.

2.17. Descripción de la tarea experimental

La tarea experimental combina un seguimiento visomotor continuo con una respuesta verbal ante eventos discretos, y se estructura del siguiente modo:

- En pantalla se presenta un estímulo visual en forma de esfera que se desplaza de manera continua por el plano bidimensional.
- El estímulo rebota al alcanzar los límites de la pantalla y modifica su dirección de forma controlada.
- A intervalos pseudoaleatorios, el estímulo cambia de color entre tres posibles valores (rojo, amarillo y verde), evitando repeticiones consecutivas. Y en intervalos de 7 a 9 segundos, se detiene durante un segundo dónde también se efectúa un cambio de color.
- El participante debe perseguir el estímulo con el cursor del ratón, simulando una tarea de persecución suave.
- Cada vez que se produce un cambio de color, el participante debe nombrar en voz alta el color mostrado, sin interrumpir el seguimiento.

Cada ensayo tiene una duración fija de 60 segundos, durante los cuales se registran de forma continua las señales necesarias para el análisis posterior.

2.18. Registro de señales visomotoras

Durante la ejecución de la tarea se registran de manera continua las posiciones del estímulo visual y del cursor del ratón. El estímulo visual es gestionado por un módulo específico que actualiza su posición en cada fotograma y registra eventos relevantes, como los cambios de color y las pausas temporales.

De forma paralela, la posición del cursor se obtiene en cada iteración del bucle principal mediante la consulta directa a la API de *Pygame*, almacenando las coordenadas junto con una marca temporal. Este procedimiento permite reconstruir la trayectoria completa del usuario y analizar su comportamiento visomotor a lo largo de la prueba.

2.19. Captura y procesamiento de la respuesta verbal

La respuesta verbal se gestiona mediante un sistema de grabación y reconocimiento de voz implementado en los módulos *whisper_loader.py* y *whisper.py*. El modelo de reconocimiento de voz

se carga en segundo plano al inicio de la aplicación mediante la clase *WhisperLoader*, evitando bloqueos en la interfaz de usuario.

Durante la ejecución de la tarea, la clase *SpeechRecorder*, definida en *whisper.py*, captura de forma continua la señal de audio durante los 60 segundos del experimento, almacenándola temporalmente en un archivo *WAV*. Una vez finalizada la prueba, el audio completo se transcribe utilizando el modelo *Whisper* previamente cargado.

El texto transcrito se normaliza eliminando acentos, caracteres no alfabéticos y diferencias de capitalización, identificando específicamente las palabras correspondientes a los colores esperados. Estas palabras se codifican numéricamente, mientras que el resto de los segmentos se etiquetan como ausencia de respuesta relevante. Dicha codificación corresponde a:

- 0 -> Sin Color. Producidos por aquellos momentos en los cuales el usuario no habla.
- 1 -> Rojo.
- 2 -> Amarillo
- 3 -> Verde

2.20. Sincronización multimodal y generación del conjunto de datos

Todas las señales registradas durante la prueba comparten una única fuente temporal basada en tiempo absoluto, lo que permite una sincronización precisa entre estímulo, cursor y respuesta verbal. Finalizada la ejecución, los registros parciales se almacenan en archivos intermedios y se fusionan posteriormente en un único conjunto de datos estructurado.

La sincronización entre señales visuales y verbales se realiza mediante un proceso de alineación temporal que asigna, a cada instante del experimento, el último evento verbal reconocido. El resultado es un fichero final en formato *Parquet* que contiene, para cada marca temporal, la posición del estímulo, la posición del cursor, el color del estímulo y el color reconocido por voz.

Una vez generado el fichero final, los archivos temporales y el audio crudo se eliminan, conservándose únicamente el conjunto de datos necesario para el análisis de los IDES.

2.21. Preparación de los datos para el análisis

Para garantizar la validez de los datos empleados en el análisis posterior, se definieron criterios prácticos de aceptación aplicables a todos los ensayos registrados. En primer lugar, se considera válido un ensayo cuya duración efectiva se sitúe en torno a los 60 segundos previstos y cuyo número de muestras sea comparable al del resto de registros, de acuerdo con el muestreo continuo de la aplicación.

Adicionalmente, se contemplan como candidatos a revisión aquellos ensayos que presenten valores extremos en métricas clave, como percentiles elevados del error de seguimiento o un rendimiento verbal claramente reducido. Estos casos no se excluyen automáticamente, sino que se marcan para su inspección mediante representaciones gráficas (*error en función del tiempo, histogramas, trayectorias estímulo–cursor, etc.*), con el fin de distinguir entre variaciones conductuales plausibles y posibles incidencias técnicas.

Los criterios definidos permiten asegurar que el conjunto de datos empleado en el cálculo de los IDES es coherente y comparable, manteniendo al mismo tiempo un control explícito sobre la calidad de los registros analizados. La Figura 4 muestra un resumen de la arquitectura llevada a cabo.



Figura 4: Diagrama general de la arquitectura llevada a cabo (Fuente: elaboración propia mediante IA generativa)

2.22. Limitaciones metodológicas

Una de las principales limitaciones del protocolo es el uso del cursor del ratón como sustituto del seguimiento ocular real. Este enfoque introduce componentes motores adicionales que pueden influir en los resultados. No obstante, permite evaluar de forma accesible la coordinación visomotora y la estabilidad del seguimiento, alineándose con el objetivo del proyecto de desarrollar indicadores digitales no invasivos y fácilmente implementables [17].

Asimismo, el estudio se limita a población sana y a un número reducido de participantes, por lo que los resultados deben interpretarse como exploratorios y servir de base para investigaciones futuras con mayor tamaño muestral y cohortes clínicas.

3

Extracción de los IDES

3.1. Introducción a los IDES

En el capítulo anterior se ha descrito el diseño experimental y el protocolo seguido para la adquisición de las señales visomotoras y verbales durante la ejecución de la tarea. A partir de dichos registros, resulta necesario definir un conjunto de métricas que permitan resumir, cuantificar y comparar de forma objetiva el desempeño de los usuarios durante los ensayos realizados.

Con este objetivo, en el presente capítulo se introducen los IDES empleados en el análisis. Estos indicadores permiten transformar señales continuas y eventos discretos en variables cuantitativas interpretables, facilitando el estudio estadístico del comportamiento observado y la comparación entre distintos ensayos y usuarios.

Como se mencionó previamente, los IDES se han seleccionado para cubrir de manera complementaria dos dimensiones principales del desempeño durante la tarea:

- El desempeño visomotor, relacionado con la precisión, estabilidad y continuidad del seguimiento del estímulo visual, y
- El desempeño cognitivo-verbal, asociado a la velocidad y exactitud de la respuesta oral ante los cambios de color del estímulo.

En total se emplearon seis IDES, de los cuales se presenta cada marco teórico, junto con la descripción de su cálculo a partir de los datos registrados en las siguientes secciones. El objetivo de

este capítulo no es interpretar los valores obtenidos ni extraer conclusiones, sino definir de manera clara y reproducible los IDES que servirán de base para el análisis de resultados desarrollado en el capítulo siguiente.

3.2. Clasificación de los IDES

Los IDES definidos en este trabajo se agrupan en función del dominio funcional que caracterizan durante la ejecución de la tarea experimental. Esta clasificación permite organizar el análisis posterior y facilita la interpretación conjunta de los resultados, sin perder de vista la naturaleza multimodal del experimento.

Dado que la tarea combina un seguimiento visomotor continuo con una respuesta verbal ante efectos discretos, los IDES se dividen en dos grandes categorías: IDES visomotoras e IDES cognitivo-verbales. Los indicadores visomotores describen la capacidad del usuario para seguir el estímulo visual de manera precisa, estable y sostenida a lo largo del tiempo.

Por otro lado, los indicadores cognitivo-verbales cuantifican la respuesta oral del estímulo ante los cambios de color del estímulo, integrando componentes de percepción visual, procesamiento cognitivo y planificación motora del habla.

En conjunto, esta clasificación establece un marco estructurado para la definición y el cálculo de los seis IDES empleados en el estudio. En los subapartados siguientes se describe de forma individual cada indicador, abordando su fundamento teórico y el procedimiento utilizado para su cálculo a partir de los datos registrados.

3.3. IDES

3.3.1. Error de seguimiento visomotor

El error de seguimiento visomotor es uno de los indicadores fundamentales para caracterizar el desempeño visomotor durante tareas de persecución continua. Este indicador cuantifica la desviación espacial entre la posición del estímulo visual y la posición del cursor controlado por el usuario, proporcionando una medida directa de la precisión con la que se mantiene el seguimiento a lo largo del tiempo.

Desde un punto de vista teórico, el error de seguimiento refleja la capacidad de integrar información visual con la respuesta motora de la mano, un proceso que involucra mecanismos de atención sostenida, coordinación ojo-mano y control motor fino.

En este trabajo, el error de seguimiento se define como la **distancia euclídea entre la posición instantánea del estímulo y la del cursor en cada instante temporal** del ensayo. A partir de esta serie temporal de errores se extraen métricas agregadas que permiten resumir el comportamiento global del usuario durante la ejecución, tales como el error medio y percentiles altos del error, que aportan información sobre desviaciones sostenidas y episodios de seguimiento deficiente.

$$error = \max(0, (\sqrt{(Xstim(t) - Xcursor(t))^2 + (Ystim(t) - Ycursor(t))^2} - R_{contacto})) \quad (1)$$

$$\text{Donde } R_{contacto} = R_{bola} + R_{tolerancia_{cursor}(\approx 38px)} \quad (2)$$

Cómo se observa en la ecuación (1), el error de seguimiento visomotor se calcula cómo el máximo de dos valores a obtener, siendo 0 en un primer instante, cuando el usuario coloca el cursor dentro del radio del estímulo, y un valor en píxeles que muestra la distancia entre el límite del radio del estímulo y el límite del radio del cursor, de ahí que se tenga en cuenta $R_{contacto}$ (2), el cual es el sumatorio de las distancias del centro al límite del radio tanto para el estímulo como para el cursor. Este indicador actúa como referencia principal del desempeño visomotor y sirve como punto de comparación para otros IDES derivados, permitiendo contextualizar el resto de métricas dentro de un marco común de precisión de seguimiento.

3.3.2. Variabilidad angular como IDE de estabilidad visomotora

La variabilidad angular es un IDE visomotor que permite evaluar la estabilidad direccional del movimiento durante el seguimiento del estímulo. A diferencia del error de seguimiento, que cuantifica la desviación espacial entre estímulo y cursor, este indicador se centra en analizar cómo de regular y consistente es la dirección del movimiento del usuario a lo largo del tiempo.

Desde un punto de vista teórico, una trayectoria de seguimiento estable se caracteriza por cambios suaves y coherentes en la dirección del movimiento, adaptándose progresivamente a la dinámica del estímulo. Por el contrario, una elevada variabilidad angular puede reflejar correcciones bruscas, movimientos erráticos o dificultades para anticipar la trayectoria, aspectos relacionados con el control motor y la planificación del movimiento en tareas de persecución continua.

En este trabajo, la variabilidad angular se calcula a partir de la dirección instantánea del vector de movimiento del cursor, obtenida a partir de las diferencias entre posiciones consecutivas. Para **cada muestra temporal se estima el ángulo de desplazamiento, y a partir de la serie temporal resultante se derivan métricas agregadas que describen la dispersión angular** del movimiento. Es decir, este IDE mide cuán recto es el seguimiento del usuario al perseguir al estímulo durante la prueba. El ángulo se obtiene calculando la hipotenusa respectiva a la trayectoria ideal que debería de haber seguido el usuario.

En particular, se emplean medidas como el valor medio absoluto del ángulo y percentiles altos de dicha distribución, que permiten resumir tanto el comportamiento típico como los episodios de mayor inestabilidad direccional. Esto, claramente, complementa al error de seguimiento visomotor, aportando información adicional sobre la calidad del movimiento, más allá de la simple distancia al estímulo. De este modo, ambos indicadores permiten caracterizar de forma conjunta precisión y estabilidad del seguimiento visomotor. [5]

3.3.3. Número de contactos del cursor sobre el estímulo.

El número de contactos es un IDE visomotor que describe la capacidad funcional del usuario para alcanzar y mantener el cursor sobre el estímulo durante la ejecución de la tarea. A diferencia de métricas puramente espaciales o direccionales, este indicador introduce un componente más discreto, basado en eventos de contacto entre el cursor y el estímulo visual.

Desde el punto de vista teórico, la ocurrencia de contactos refleja la coordinación ojo–mano y la capacidad de sincronizar el movimiento del cursor con la trayectoria del estímulo, siendo una tarea de persecución en la que el estímulo, dado cierto tiempo, se detiene, la aparición de contactos es un factor clave, más aún, habiendo presentes cambios tanto direccionales como de velocidad.

En este trabajo, se considera un impacto o alcance cuando la **distancia entre el cursor y el centro del estímulo se sitúa por debajo de un umbral definido por el radio del estímulo**. En otras palabras, cuando el radio del cursor coincide o se superpone sobre el del estímulo, contabilizándose así, un número discreto.

El valor final del indicador se obtiene como el número total de contactos registrados durante el ensayo, y puede analizarse de forma conjunta con otros IDEs visomotoras para caracterizar tanto la precisión como la continuidad del seguimiento. Este indicador aporta una visión funcional del desempeño, al reflejar no solo la proximidad al estímulo, sino la consecución efectiva del objetivo durante la tarea.

3.3.4. Tiempo efectivo en objetivo

El tiempo efectivo en objetivo es un indicador de desempeño visomotor que cuantifica la capacidad del usuario para mantener el cursor sobre el estímulo de forma sostenida durante la ejecución de la tarea. A diferencia del número de contactos, que contabiliza eventos discretos de contacto, este indicador incorpora explícitamente la dimensión temporal del seguimiento.

Desde un punto de vista teórico, la permanencia del cursor sobre el estímulo refleja no solo la precisión espacial, sino también la estabilidad del control motor a lo largo del tiempo. En tareas de seguimiento continuo, la capacidad de mantener el contacto durante intervalos prolongados es indicativa de un seguimiento fluido y coordinado, mientras que contactos breves o intermitentes pueden asociarse a correcciones frecuentes o pérdida de estabilidad.

En este trabajo, el tiempo efectivo en objetivo se define como el **tiempo total acumulado durante el cual el cursor permanece dentro del área del estímulo**, compenetrándose así con el número de contactos. A partir de la serie temporal de posiciones sincronizadas, se identifican los intervalos consecutivos en los que el cursor se mantiene dentro del radio del estímulo y se calcula su duración. El valor del indicador se obtiene como la suma de estos intervalos de contacto, pudiendo expresarse tanto en términos absolutos como relativos al tiempo total del ensayo, permitiendo diferenciar entre contactos puntuales o "choques" y un control sostenido del cursor sobre el estímulo.

3.3.5. Latencia verbal ante el cambio de color del estímulo

La latencia de respuesta verbal es un indicador de desempeño cognitivo-verbal que cuantifica el tiempo transcurrido entre un cambio de color del estímulo y la emisión de la respuesta oral correspondiente por parte del usuario, como se muestra en la ecuación (3). Este indicador introduce una dimensión temporal ligada al procesamiento cognitivo y a la planificación de la respuesta verbal bajo condiciones de doble tarea.

Desde el punto de vista teórico, la latencia verbal integra varios procesos encadenados: detección del cambio visual, identificación del color, selección de la respuesta lingüística adecuada y ejecución motora del habla.

$$latencia = t_{habla} - t_{color_mostrado} \quad (3)$$

En este trabajo, la latencia de respuesta verbal se calcula a partir de la **diferencia temporal entre el instante del cambio de color del estímulo y el instante asociado a la respuesta verbal reconocida**. Los eventos de cambio de color y las respuestas verbales se encuentran sincronizados mediante un mismo eje temporal, lo que permite estimar la latencia para cada evento de forma consistente.

A partir de las latencias individuales se obtienen métricas agregadas, como el valor medio de latencia por ensayo o su distribución por color, que permiten resumir el comportamiento global del usuario, e identificar patrones asociados a pronunciar una palabra u otra.

3.3.6. Precisión verbal

La precisión verbal es un indicador de desempeño cognitivo-verbal que evalúa la exactitud de la respuesta oral del usuario ante los cambios de color del estímulo. A diferencia de la latencia de respuesta, que se centra en la dimensión temporal, este indicador cuantifica el grado de corrección de la respuesta emitida, incorporando un componente semántico y de control de la respuesta.

Desde un punto de vista teórico, la precisión verbal refleja la capacidad de identificar correctamente el estímulo visual y seleccionar la respuesta lingüística adecuada bajo condiciones de doble tarea. En paradigmas que combinan seguimiento visomotor y denominación verbal, la tasa de aciertos y errores proporciona información complementaria sobre el desempeño cognitivo, especialmente en contextos donde la atención y el procesamiento simultáneo pueden verse comprometidos.

En este trabajo, la precisión verbal se calcula a partir de la **comparación entre el color del estímulo presentado y el color reconocido a partir de la respuesta verbal**. Cada evento es catalogado como acierto o error, obteniendo nuevas métricas como el ratio de acierto o error, o la propia distribución de respuesta por color, tomando valores de 0 a 1.

En este capítulo se han definido y justificado los seis IDES empleados en el estudio, describiendo su fundamento teórico y el procedimiento utilizado para su cálculo a partir de los datos registrados. Estos IDES permiten caracterizar el desempeño del usuario desde una perspectiva multimodal, integrando métricas visomotoras y cognitivo-verbales.

En el capítulo siguiente se presentan los resultados obtenidos a partir de la aplicación de estos indicadores al conjunto de ensayos experimentales, junto con el análisis estadístico correspondiente.

4

Resultados Obtenidos

4.1. Descripción del conjunto de datos

En este apartado se describen las características del conjunto de datos generado tras la aplicación del protocolo experimental descrito en el Capítulo 9. El estudio comprende un total de 20 ensayos consecutivos, cada uno con una duración de 60 segundos, realizados bajo condiciones experimentales homogéneas.

Cada ensayo da lugar a un fichero en formato *Parquet* que contiene las señales sincronizadas necesarias para el análisis posterior de los IDES. La Tabla 1 resume los ensayos registrados, incluyendo su número de muestras y archivo asociado.

| USER_ID | Nº MUESTRAS | ARCHIVO PARQUET |
|---------|-------------|-------------------|
| 1 | 3444 | final_001.parquet |
| 2 | 3555 | final_002.parquet |
| 3 | 3562 | final_003.parquet |
| 4 | 3486 | final_004.parquet |
| 5 | 3476 | final_005.parquet |
| 6 | 3525 | final_006.parquet |
| 7 | 3587 | final_007.parquet |

| | | |
|----|------|-------------------|
| 8 | 3559 | final_008.parquet |
| 9 | 3526 | final_009.parquet |
| 10 | 3452 | final_010.parquet |
| 11 | 3540 | final_011.parquet |
| 12 | 3602 | final_012.parquet |
| 13 | 3600 | final_013.parquet |
| 14 | 3590 | final_014.parquet |
| 15 | 3609 | final_015.parquet |
| 16 | 3570 | final_016.parquet |
| 17 | 3629 | final_017.parquet |
| 18 | 3585 | final_018.parquet |
| 19 | 3592 | final_019.parquet |
| 20 | 3606 | final_020.parquet |

Tabla 1: Lista de ensayos experimentales registrados

4.2. Flujo de procesamiento y extracción de IDES (resumen operativo)

El procesamiento de los datos sigue el flujo descrito en el capítulo de Metodología y protocolo experimental, y se resume esquemáticamente en la Figura 6. A partir de los ficheros Parquet generados por cada ensayo, se procede a la depuración, sincronización estímulo–respuesta y cálculo de métricas espaciales y temporales necesarias para la obtención de los IDES.



Figura 5: Etapas del proceso de extracción y análisis de datos (Fuente: elaboración propia mediante IA generativa)

4.3. Comparabilidad entre ensayos

Antes de comparar los IDES, se debe explicar la validación de las pruebas tomadas en los usuarios, tanto aspectos como integridad de datos como consistencia de estos mismos.

Para empezar, se evaluaron métricas para la consistencia de cada ensayo, en este caso se aseguró que tanto la duración efectiva de cada prueba como el número de muestras estuvieran en rangos estrechos de diferencia. El motivo principal de esto viene de que todas las pruebas corresponden a ejecuciones de una misma tarea, en este caso, la API desarrollada de *Python*, por lo que el número de muestras deben estar alrededor de 3500 y la duración por cada prueba, 60s, exceptuando algunas milésimas, ya que al cargar el modelo *Whisper*, éste toma un lapsus de tiempo indefinido breve para funcionar. Todo esto se puede ver reflejado en la Tabla 2.

| VARIABLE | MEDIA | MÍN | MÁX |
|--|---------|----------|---------|
| Duración del ensayo, <i>duration_s</i> (s) | 60,0082 | 59.99598 | 60,0153 |
| Nº de muestras, <i>n_samples</i> | 3554.75 | 3444 | 3629 |

Tabla 2: Criterios de aceptación de los ensayos

A continuación, se aplicaron criterios prácticos de aceptación para identificar ensayos potencialmente atípicos. Se consideraron candidatos a inspección aquellos registros que presentaban valores extremos en métricas de seguimiento (*por ejemplo, percentiles elevados del error, como p95/p99*) o un rendimiento verbal marcadamente bajo (tasa de acierto o tasa de respuesta en ventana reducida).

En el caso de este trabajo, dicho percentil se tiene en cuenta para los IDES 1, 2 y 5 con el fin de aportar un estudio estadístico más robusto pese al pequeño número de ensayos realizados. La media o la mediana describen el rendimiento típico, pero pueden ocultar episodios breves de degradación, ya sea por pérdidas puntuales de seguimiento, lapsos atencionales o demoras puntuales en la respuesta verbal, que, aunque representen una fracción pequeña del ensayo, han sido considerados suficientemente relevantes como para tenerlos en cuenta.

El p95 permite caracterizar la "cola" de la distribución sin depender exclusivamente de un único valor: identifica el umbral por encima del cual se sitúa el 5% de los valores más altos del ensayo, reflejando

así los momentos de peor desempeño de una manera más estable, ofreciendo así un compromiso entre sensibilidad a episodios críticos y robustez estadística, minimizando el impacto de outliers aislados y favoreciendo la interpretabilidad de los resultados, por lo que siempre se observarán los peores casos para cada ensayo realizado por cada usuario.

El resultado de este filtrado fue el que se muestra en la Tabla 3. Como se puede observar, 4 fueron excluidas por integridad del seguimiento, la causa principal fue la pérdida en la continuidad del intento de mantener la persecución suave buscada en este trabajo, debido a un malentendido en las instrucciones mostradas por pantalla durante la prueba. En cuanto al ensayo excluido por audio/voz no válida, este fue atribuido al sujeto de 48 años el cual olvidó por completo el recitar los colores en voz alta.

| Total de Pruebas Registradas | Excluidas por integridad de seguimiento | Excluidas por audio/voz no válida | Excluidas por duración/ventanas no válidas | Pruebas Analizadas |
|------------------------------|---|-----------------------------------|--|--------------------|
| 25 | 4 | 1 | 0 | 20 |

Tabla 3: Flujo de inclusión/exclusión de pruebas

Como apoyo a la inspección de ensayos marcados, se emplearon visualizaciones diagnósticas (Figuras 6 y 7), útiles para distinguir valores extremos atribuibles al comportamiento del participante de posibles incidencias técnicas, previas al análisis general para cada IDE de todos los usuarios. Dichas gráficas se obtuvieron mediante el *script* de *Jupyter Notebook*, en el cual aparte de obtener el archivo *Excel* general correspondiente a la prueba, también se muestran gráficas individuales para cada IDE, las cuáles sirven para tener una idea de cómo interpretar los datos desde un punto de vista general.

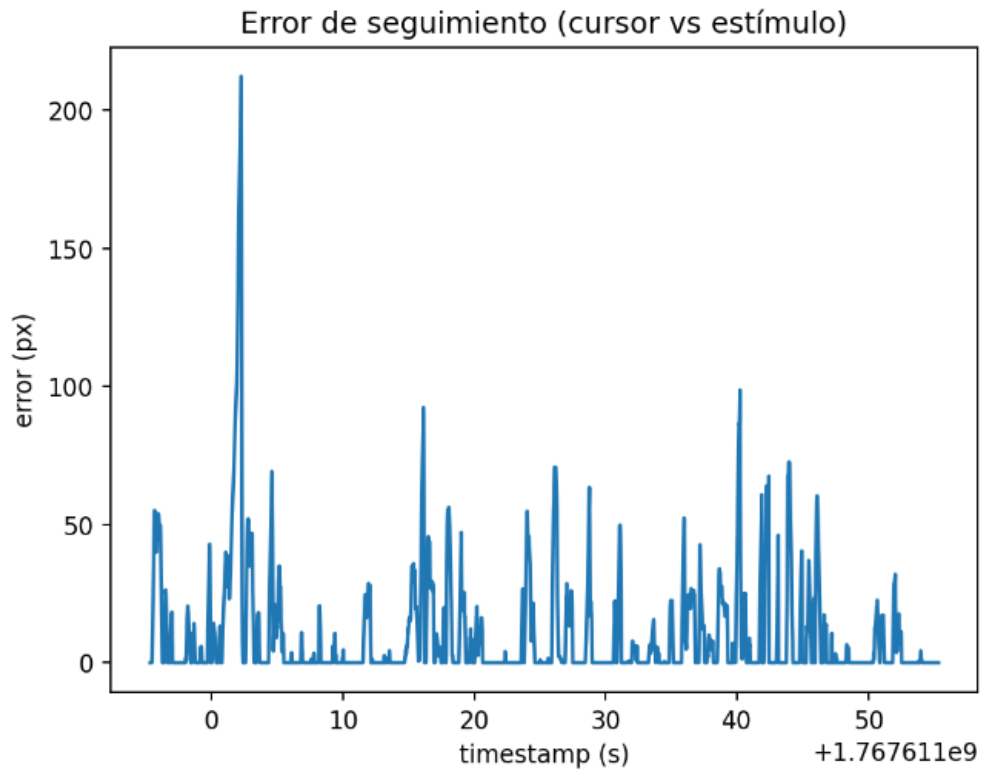


Figura 6: Error de seguimiento visomotor en función del tiempo (usuario 1)

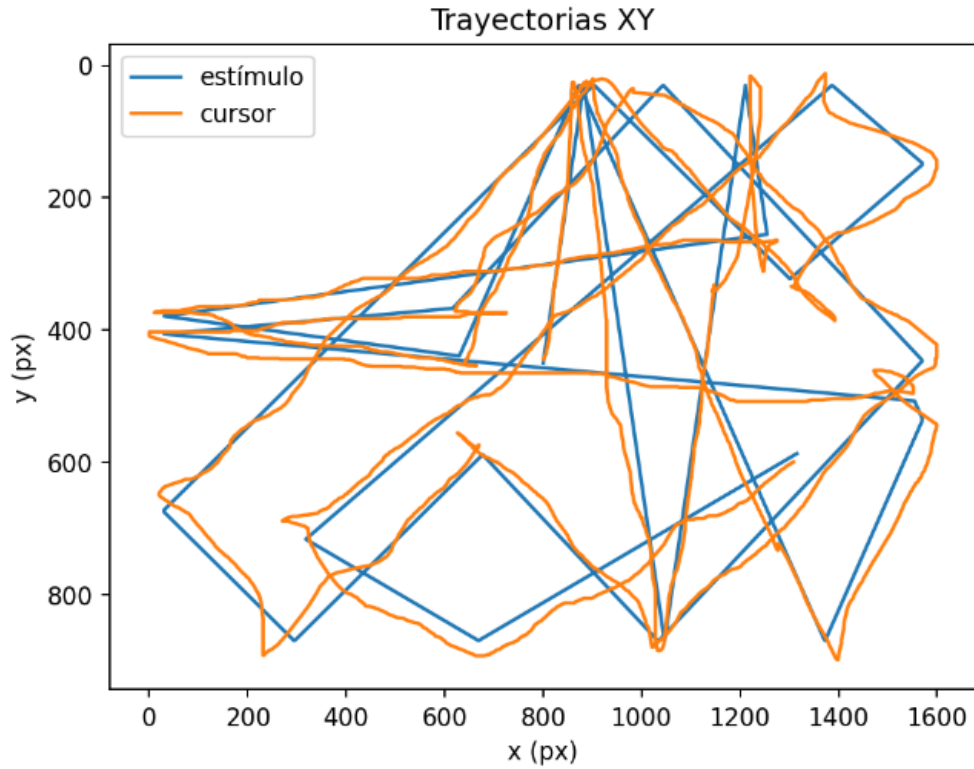


Figura 7: Trayectorias del estímulo y del cursor durante el ensayo (usuario 1)

En conjunto, este control de calidad respalda que los ensayos analizados son suficientemente homogéneos para comparar IDES entre usuarios, y que los casos extremos han sido identificados y verificados. En los apartados siguientes se prioriza el análisis agregado (tendencias y dispersión) y, cuando resulta necesario para justificar valores más extremos.

4.4. Visión global de los usuarios

En este apartado, con el fin de complementar el posterior análisis individual de cada IDE, y tener una visión de conjunto previa, se presenta en la Figura 8 una tabla comparativa de métricas por usuario. En ella se incluyen métricas representativas del seguimiento, interacción, estabilidad angular y rendimiento verbal por parte de los usuarios.

Dado que cada métrica tiene su propia unidad y escala, cada columna se transforma a *z-score*, calculado como se indica en la ecuación (4):

$$Z = \frac{x - \mu}{\sigma} \quad (4)$$

donde x es el valor del usuario en esa métrica, y μ y σ son la media y desviación estándar del conjunto analizado para dicha métrica. Esta transformación no pretende comparar magnitudes absolutas entre las métricas, sino, expresar los valores positivos (encima de la media del conjunto), y los negativos (por debajo), estando estos recortados entre $[-2,2]$

Además, facilitando la lectura global, los usuarios están ordenados por atipicidad (5), es decir, de arriba abajo según la media del valor absoluto de sus correspondientes *z-scores*. Esto se puede ver reflejado en la siguiente fórmula:

$$\textit{Atipicidad} = \textit{mean} (|z_1|, |z_2|, \dots, |z_n|) \quad (5)$$

De esta manera, la Figura 8 permite identificar rápidamente qué usuarios se alejan más del comportamiento medio cuando se consideran varias dimensiones a la vez. En la práctica, aquellos en posiciones superiores de la tabla, presentan más desviaciones marcadas en sus métricas que los que se encuentran más abajo, siendo estos últimos perfiles con valores propios de la media del conjunto y, cómo se observa, en este caso, la atipicidad de estos no depende explícitamente de la edad.

| | error_p95_px | error_mean_px | on_target_fraction | touches_count | mean_contact_duration_s | ang_p95_abs_deg | ang_prop_gt10deg | accuracy_voice_nonzero | latency_p95_s | latency_success_rate_within_window |
|--------------|--------------|---------------|--------------------|---------------|-------------------------|-----------------|------------------|------------------------|---------------|------------------------------------|
| USER 14 (60) | 1.54 | 1.85 | -1.29 | -1.34 | -1.29 | 1.77 | 2.00 | 0.63 | -0.38 | 0.21 |
| USER 12 (43) | 0.89 | 1.14 | -1.25 | -1.39 | -1.11 | 1.30 | 1.17 | -2.00 | -0.57 | 0.96 |
| USER 13 (66) | 2.00 | 1.62 | -1.20 | -0.76 | -1.58 | -0.19 | 0.36 | 1.46 | 1.57 | 0.97 |
| USER 19 (55) | 1.43 | 1.55 | -1.17 | -1.08 | -1.24 | 2.00 | -0.21 | 0.63 | 0.12 | 1.05 |
| USER 10 (26) | -0.83 | -0.86 | 1.00 | 0.92 | 0.87 | -0.83 | 0.01 | -1.73 | 2.00 | -0.43 |
| USER 20 (48) | 1.12 | 1.36 | -1.26 | -1.71 | -0.63 | 0.51 | 1.25 | -0.44 | -1.00 | 0.60 |
| USER 9 (18) | -0.90 | -0.91 | 1.23 | 0.39 | 1.83 | -0.96 | -1.21 | 0.45 | -0.42 | -0.78 |
| USER 11 (16) | -0.91 | -0.88 | 0.82 | 0.50 | 1.07 | -0.96 | -0.56 | 0.96 | -1.16 | 0.74 |
| USER 6 (20) | -0.87 | -0.87 | 1.29 | 0.92 | 1.27 | -0.88 | -0.48 | 0.48 | 0.16 | -1.17 |
| USER 8 (22) | -0.86 | -0.87 | 0.94 | 1.08 | 0.60 | -0.96 | -1.40 | 0.08 | -0.28 | -1.33 |
| USER 15 (40) | 0.39 | 0.66 | -1.05 | -1.39 | -0.39 | 0.85 | 1.37 | -0.69 | 0.02 | 1.47 |
| USER 3 (28) | -0.82 | -0.84 | 0.87 | 0.45 | 1.20 | -1.04 | -1.21 | 0.50 | 0.11 | -0.92 |
| USER 18 (50) | 1.15 | 0.80 | -1.04 | -0.92 | -0.99 | -0.10 | -1.26 | -0.63 | -0.69 | 0.32 |
| USER 5 (19) | -0.89 | -0.89 | 0.93 | 1.08 | 0.60 | -0.80 | -0.87 | 0.03 | 0.47 | -1.13 |
| USER 17 (36) | -0.27 | -0.20 | -0.78 | -0.45 | -0.83 | 0.25 | 0.06 | -0.69 | -1.60 | 2.00 |
| USER 7 (24) | -0.86 | -0.89 | 1.04 | 0.97 | 0.86 | -0.71 | 0.34 | -0.26 | -0.10 | -0.78 |
| USER 1 (32) | -0.74 | -0.77 | 0.57 | 1.44 | -0.15 | -0.54 | -0.65 | 0.02 | 0.71 | -0.98 |
| USER 2 (30) | -0.76 | -0.66 | 0.64 | 0.97 | 0.31 | -0.78 | 0.31 | -0.49 | -0.23 | -1.06 |
| USER 16 (38) | 0.33 | 0.00 | -0.60 | 0.03 | -0.88 | 0.90 | -0.46 | 1.78 | -0.57 | -0.33 |
| USER 4 (34) | -0.29 | -0.35 | 0.32 | 0.29 | 0.46 | 0.89 | 1.26 | 0.58 | 0.74 | 0.37 |

Figura 8: Tabla comparativa en z-score por usuario, ordenada por índice de atipicidad (media de |z|). Cada celda muestra la desviación relativa de una métrica respecto a la media del conjunto.

Por último, como ya se mencionó, esta figura indica una contextualización previa al análisis de IDE por IDE, sin inferir patrones clínicos ni sustituir la interpretación detallada de cada indicador. Tiene dos funciones, ofrecer una panorámica global del conjunto analizado, y orientar localización de valores extremos o perfiles diferenciados que, cuando es necesario, se contrastan con las visualizaciones por prueba mostradas a continuación y así justificar el origen de las desviaciones sin sobreinterpretar dichos resultados.

4.5. IDE 1: Error de Seguimiento Visomotor

El error de seguimiento visomotor cuantifica la desviación del cursor respecto al estímulo fuera del radio de contacto: a mayor desviación más error. En este proyecto se deriva un vector de error en el tiempo como $error(t) = \max(0, d(t) - R_{contacto}) (\delta)$, siendo $d(t)$ la distancia euclídea cursor-estímulo. En este proyecto se emula conceptualmente una tarea de persecución suave mediante el cursor, registrando en cada instante la posición del estímulo y del cursor para derivar un vector de error a lo largo del tiempo.

Para resumir este IDE se emplearon métricas descriptivas como el error medio y percentiles altos ($p95/p99$), la justificación de su uso fue previamente explicada en el apartado 11.4.

La Tabla 4 resume el error de seguimiento por ensayo. Se reporta el error medio y percentiles altos ($p95$ y $p99$) para capturar no solo el rendimiento típico, sino también episodios de pérdida de seguimiento que afectan a la estabilidad del control visomotor.

| User_Id | error_mean_px | error_p95_px | error_p99_px |
|---------|---------------|--------------|--------------|
| 1 | 9,900531 | 47,36197 | 78,80226 |
| 2 | 14,01365 | 44,46361 | 307,7342 |
| 3 | 7,518377 | 38,81407 | 58,20242 |
| 4 | 25,48399 | 94,01286 | 416,3907 |
| 5 | 5,772673 | 31,38703 | 64,10228 |
| 6 | 6,227097 | 33,31334 | 93,55554 |
| 7 | 5,58014 | 34,95271 | 56,17919 |
| 8 | 6,453894 | 34,15755 | 60,15157 |
| 9 | 4,822877 | 29,8127 | 55,66833 |
| 10 | 6,763909 | 37,46808 | 65,09102 |
| 11 | 6,057343 | 29,39312 | 49,87589 |
| 12 | 79,88102 | 218,3993 | 277,4628 |
| 13 | 97,47701 | 350,981 | 491,195 |
| 14 | 106,0859 | 286,616 | 356,7707 |
| 15 | 62,48631 | 165,6889 | 223,1846 |
| 16 | 38,39801 | 159,5177 | 217,3845 |
| 17 | 31,08447 | 96,26401 | 202,8471 |
| 18 | 67,54746 | 245,5564 | 323,5618 |
| 19 | 95,09001 | 274,4037 | 333,1041 |
| 20 | 88,23877 | 242,0514 | 324,7318 |

Tabla 4: Resumen de métricas de seguimiento visomotor para los 20 ensayos

En la Figura 9 se ve que, a medida que la edad aumenta, el error de seguimiento en el percentil 95 incrementa de forma muy marcada. Las medianas por cada "boxplot" correspondiente a cada grupo son las siguientes:

- 16-25 años: 33.35 px
- 26-35 años: 44.46 px
- 36-45 años: 162.60 px
- 46-55 años: 245.56 px
- 56-70 años: 318.80 px

Observando estos valores, se aprecia una tendencia ascendente, de donde se puede destacar que el grupo más joven (16-25 años) con una mediana de 32 px apenas, presenta una gran diferencia respecto al de mayor edad (56-70 años) con una mediana aproximada de 319 px, lo que equivale a un error 10 veces mayor en el peor 5% de los casos. Teniendo en cuenta que los saltos tampoco son uniformes, ya que, el incremento de 16-25 a 26-35 años es de +12 px, y por otro lado de 26-35 a 36-45 años es de +118px, lo que supone un aumento de incremento 10 veces mayor que el primer salto. Esto solo confirma que el punto de inflexión es mucho más marcado a partir de los 36 años y el peor desempeño individual empeora de manera significativa a medida que avanza la edad.

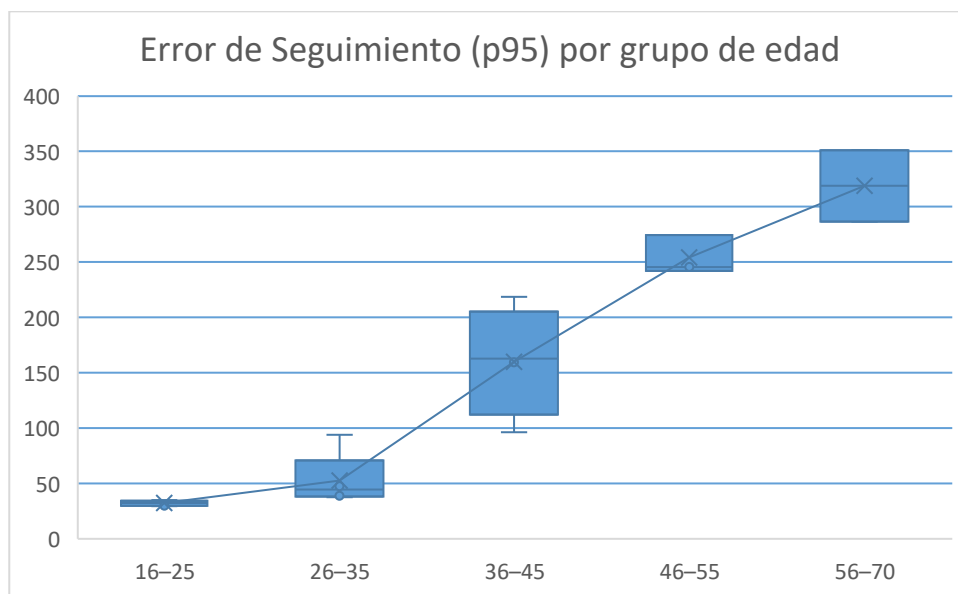


Figura 9: Error de seguimiento (p95) por grupo de edad

Por lo que el error de seguimiento, complementado por los percentiles previamente mencionados, se comporta como un candidato robusto a indicador de desempeño digital porque es sensible a diferencias interejecución, es decir, permite detectar pérdidas puntuales de control frente a deterioro sostenido, y muestra una tendencia clara por grupos etarios en el *dataset*, lo que respalda su uso como variable base para comparaciones futuras en cohortes clínicas [8].

4.6. IDE 2: Variabilidad angular (estabilidad direccional del seguimiento)

La variabilidad angular evalúa la coherencia direccional entre el movimiento del estímulo y el movimiento del cursor. Mientras que el IDE 1 cuantifica "cuán lejos" está el cursor del estímulo, este IDE cuantifica "cuán alineada" es la dirección de persecución. Por ello, es especialmente útil para detectar seguimientos erráticos que pueden no reflejarse completamente en la distancia media.

Para resumir este IDE se emplearon métricas sobre el ángulo absoluto (*en grados*): media y percentil 95 del ángulo absoluto, así como una medida de dispersión (*desviación estándar*) y dos proporciones de episodios de desalineación (*proporción de muestras con $|\text{ángulo}| > 5^\circ$ y $|\text{ángulo}| > 10^\circ$*). En conjunto estas medidas permiten caracterizar tanto el rendimiento típico como la frecuencia de eventos de seguimiento direccionalmente inestable, aunque finalmente el análisis se ha realizado utilizando la medida para el percentil 95.

En la Tabla 5 se observa que algunos registros presentan valores claramente más altos (*p.ej medias absolutas y p95 elevados*), lo que sugiere que el cursor no solo se aleja, sino que además no acompaña la dirección del estímulo durante una fracción relevante del ensayo. Esto es coherente con lo que ya explicas en el marco teórico: el seguimiento suave "ideal" requiere predicción y control continuo; cuando fallan, aparecen correcciones direccionales y desalineación.

| user_id | ang_mean_abs_deg | ang_std_deg | ang_p95_abs_deg | ang_prop_gt5deg | ang_prop_gt10deg |
|---------|------------------|-------------|-----------------|-----------------|------------------|
| 1 | 15,21425 | 28,66218 | 53,60235 | 0,689769 | 0,432343 |
| 2 | 15,54464 | 25,1818 | 45,11737 | 0,721617 | 0,530404 |
| 3 | 12,74127 | 23,50244 | 35,45226 | 0,613512 | 0,375857 |
| 4 | 23,93529 | 39,70789 | 105,7768 | 0,774364 | 0,626412 |
| 5 | 13,74004 | 26,15738 | 44,37808 | 0,641087 | 0,41053 |
| 6 | 14,20532 | 22,98983 | 41,49767 | 0,724431 | 0,449541 |
| 7 | 15,07999 | 26,08036 | 47,46649 | 0,724589 | 0,533378 |
| 8 | 13,72165 | 26,16827 | 38,46735 | 0,75529 | 0,356314 |
| 9 | 13,11072 | 23,70116 | 38,38215 | 0,714528 | 0,375764 |
| 10 | 14,35775 | 25,79197 | 43,31556 | 0,662886 | 0,499823 |
| 11 | 13,5352 | 22,63961 | 38,32633 | 0,755219 | 0,442088 |
| 12 | 28,64101 | 47,54752 | 120,6247 | 0,81578 | 0,617083 |
| 13 | 18,3438 | 29,63659 | 66,26279 | 0,762266 | 0,534902 |
| 14 | 31,18683 | 48,74185 | 137,5395 | 0,845593 | 0,721721 |
| 15 | 28,2069 | 43,59403 | 104,3074 | 0,771359 | 0,638187 |
| 16 | 20,37605 | 38,28703 | 106,1551 | 0,802346 | 0,452186 |
| 17 | 18,09509 | 34,04192 | 82,4227 | 0,764203 | 0,504449 |
| 18 | 15,19658 | 31,09957 | 69,69703 | 0,567959 | 0,371149 |
| 19 | 30,11942 | 55,58831 | 155,4603 | 0,688247 | 0,477457 |
| 20 | 26,44418 | 39,90721 | 91,92328 | 0,860278 | 0,625457 |

Tabla 5: Variables empleadas en el cálculo del indicador IDE 2

En enfermedades neurodegenerativas, este tipo de inestabilidad se asocia a disfunción de circuitos frontoparietales (*control atencional/ejecutivo*), cerebelo (*ajuste fino del movimiento*) y redes oculomotoras (*persecución suave*). En este caso, aunque el seguimiento es realizado por ratón, el IDE mantiene valor como marcador digital accesible de estabilidad direccional [8].

Para contextualizar la dispersión por edad de la estabilidad direccional, la Figura 10 representa el percentil 95 del ángulo absoluto por grupo de edad. Esta figura permite comprobar si los valores angulares extremos se agrupan en determinados rangos de edad o si corresponden a ejecuciones puntuales.

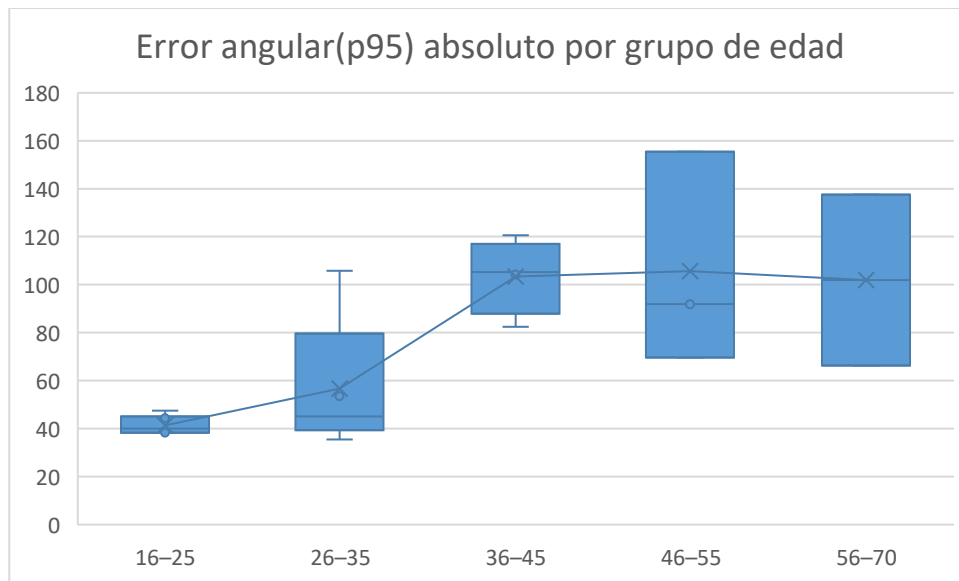


Figura 10: Métrica angular (p95 abs) por grupo de edad

Como se puede observar en dicha figura, la tendencia general muestra un aumento de la variabilidad angular hasta los 35 años, con valores en torno a 40-45° en la mediana de dicho "boxplot", lo que nos lleva a una baja dispersión y un seguimiento bastante estable incluso en el peor % de los casos. A partir de los 35 años, hasta los 45, se produce un aumento abrupto de la mediana, unos 60° aproximadamente, donde todo el grupo se desplaza a valores mayores de 80°, mostrando así que no es un "outlier" sino un cambio global del grupo. Por último, para los dos últimos grupos etarios, la mediana se sigue manteniendo igual de elevada (92-105° apróx.), pero en estos casos la dispersión de los datos aumenta considerablemente y aparecen valores extremos muy altos (hasta 155°).

Entendiendo esto, cabe decir que en jóvenes se sigue manteniendo un patrón estable, cómo también mostró el IDE 1, hay un punto de inflexión a partir de los 36 años, dando lugar a una alta variabilidad angular en edades posteriores y una mayor aparición de episodios extremos. No es un crecimiento lineal perfecto, sino un cambio brusco seguido de estabilización en valores altos con mayor dispersión por grupo etario.

Las Figuras 11 y 13 muestran un comportamiento dentro del conjunto del *dataset* bastante estable. Éstas corresponden al Usuario 1(32 años) y reflejan una variabilidad angular prácticamente inexistente y uniforme. En la serie temporal hay una banda estrecha sin apenas oscilaciones en el eje Y, lo que resalta un seguimiento angular estable, y, en cuanto al histograma, la mayoría de los ángulos están reunidos entre aproximadamente -50° y 50°, confirmando una dispersión reducida. Por otro lado, las Figuras 12 y 14, correspondientes al Usuario 14(66 años), muestran picos y saltos marcados a lo largo del tiempo. En el histograma, los ángulos se distribuyen de forma mucho

más amplia, abarcando casi todo el rango de -150° a 150° , lo que evidencia una variabilidad mucho mayor.

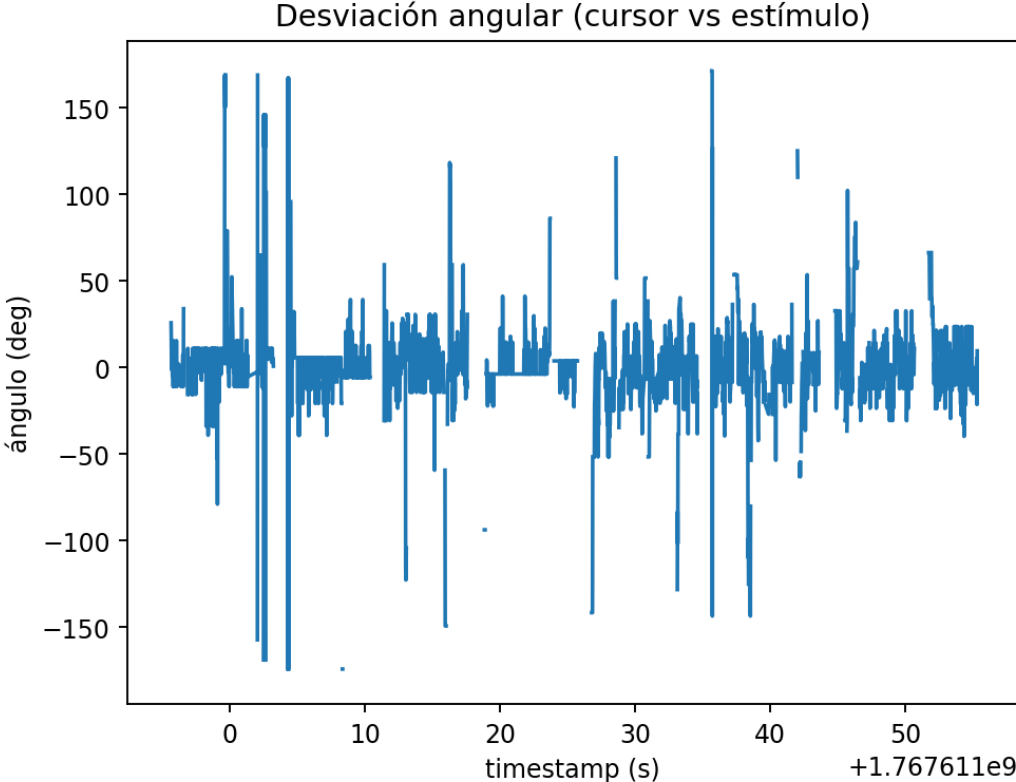


Figura 11: Serie temporal del ángulo en función del tiempo (ejecución con seguimiento estable)

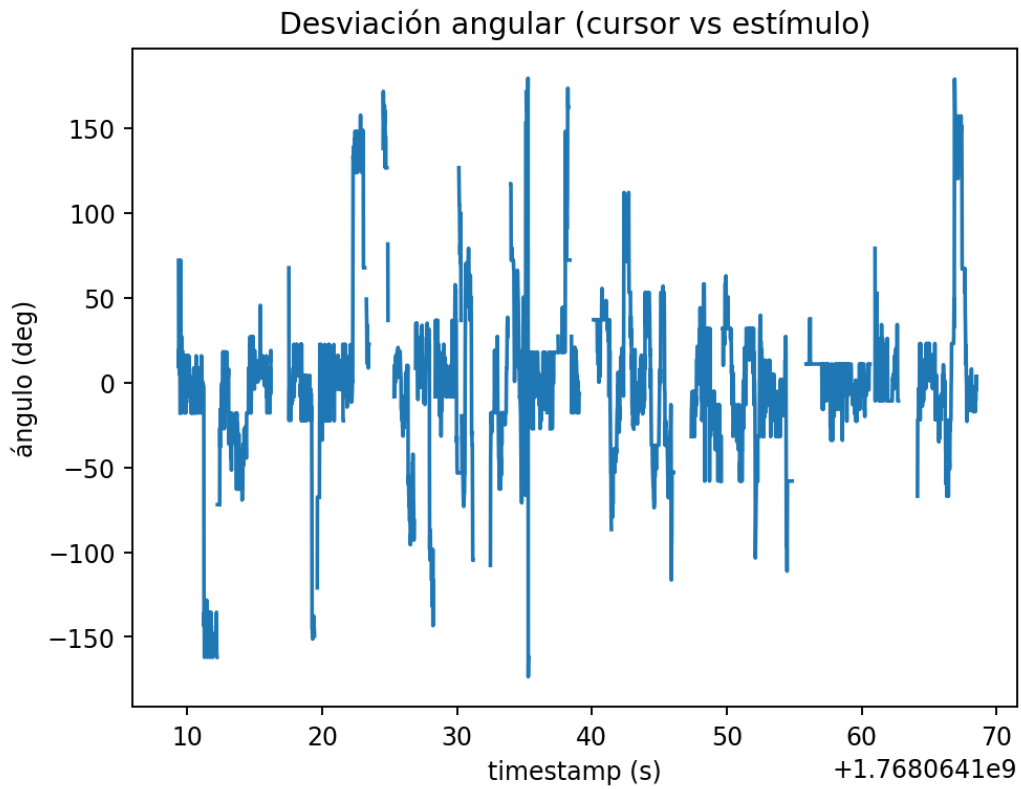


Figura 12: Serie temporal del ángulo en función del tiempo (ejecución con seguimiento inestable)

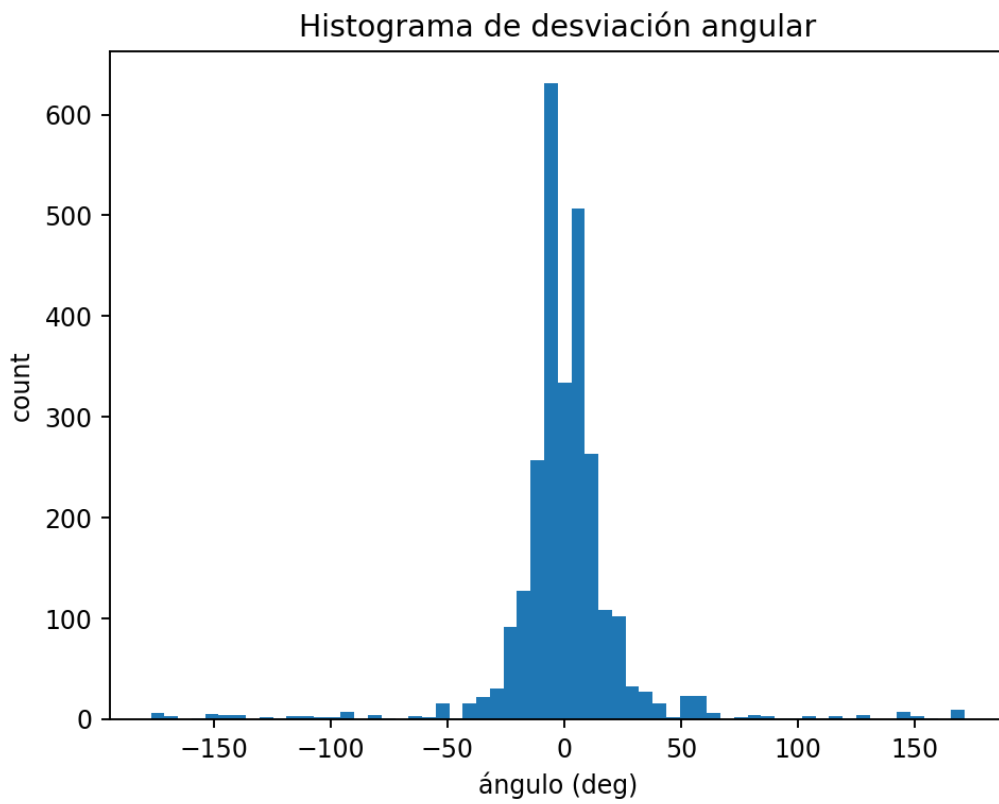


Figura 13: Histograma del ángulo durante una ejecución con seguimiento estable

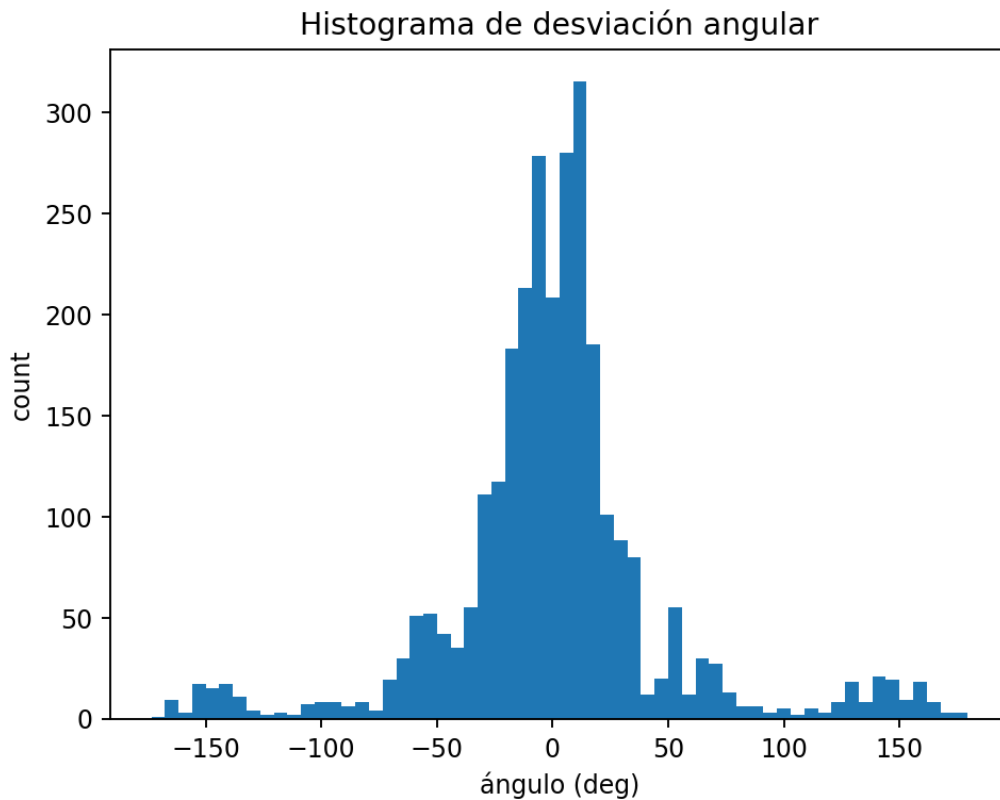


Figura 14: Histograma del ángulo durante una ejecución con seguimiento inestable

En cuanto a la Figura 15, en este caso, muestra claramente una relación positiva entre ambas métricas en su percentil 95, conforme aumenta el error de seguimiento también lo hace la variabilidad angular. Los usuarios de 16 a 32 años se concentran en la zona izquierda inferior del gráfico ($[38-50]^o$ y $[29-50]px$), siendo ésta una agrupación compacta y homogénea, lo que indica que en usuarios en este rango etario se mantienen un seguimiento estable, baja variabilidad angular y una consistencia interindividual. A continuación, en cuanto se alcanza al usuario de 34 años, se produce un punto de transición donde la distribución se caracteriza por aumentos en la variabilidad angular y el error de seguimiento, aumentando la dispersión, lo que refuerza lo que ya se observaba en los "boxplots": hay un punto de inflexión notorio en el grupo etario de 36-45 años. Teniendo esto en cuenta, cabe destacar el caso del usuario de 66 años, que pese a tener un error de seguimiento muy alto, su variabilidad angular no es destacable en términos de percentil 95, situándose incluso en valores aproximados a los observados en el rango etario de 16-35 años. No obstante, este resultado no debe interpretarse de forma aislada. Tal y como se mostró en las Figuras 12 y 14, donde se analiza su evolución angular en crudo, el desempeño global no fue óptimo. En consecuencia, el valor del percentil 95 no refleja completamente el patrón real de seguimiento observado en la serie temporal, por lo que este caso debe considerarse como una situación particular dentro del análisis general.

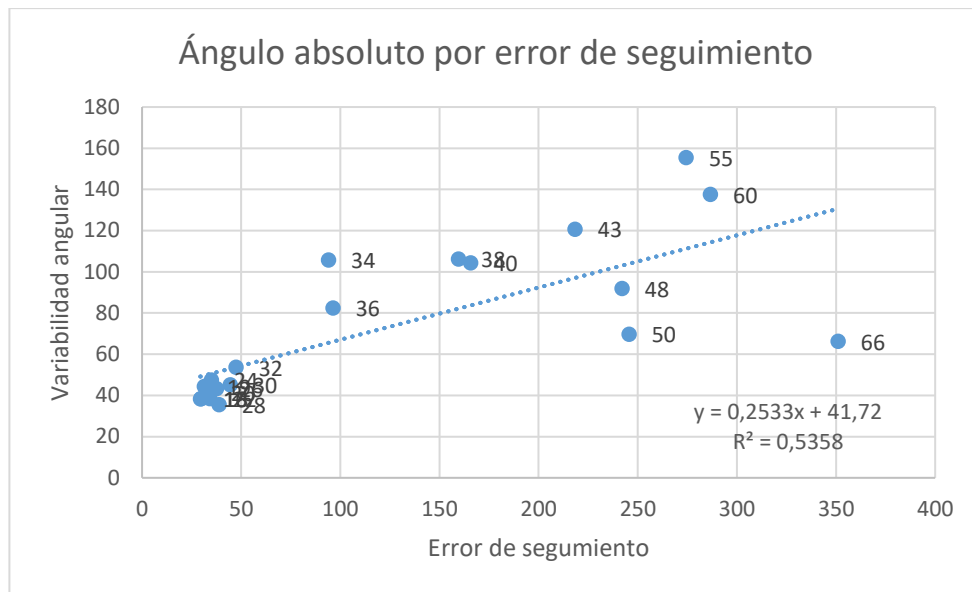


Figura 15: Relación entre el percentil 95 del ángulo absoluto y el percentil 95 del error de seguimiento

4.7. IDE 3: Contactos del cursor sobre el estímulo (alcances exitosos)

El número de contactos cuantifica cuántas veces el cursor consigue colocarse sobre el estímulo durante la prueba. Este IDE resume la coordinación ojo-mano, la precisión de control motor fino y la capacidad de mantener la persecución efectiva en una tarea continua. Todo "contacto" se define dentro de este TFG como el hecho de entrar y salir de la superficie del estímulo en cualquier momento durante la duración del ensayo.

En este trabajo se reporta el número total de contactos durante los 60s (*touches_count*). Adicionalmente, para aportar una lectura más robusta del "contacto efectivo", se complementa con métricas temporales del contacto descritas en el IDE siguiente.

En la tabla 6 hay una variabilidad notable entre registros, con valores altos en algunos casos y descensos marcados en otros. Eso es exactamente lo que se espera en un indicador de coordinación: cuando el control visomotor es robusto, el usuario recupera el contacto muchas veces; cuando hay una mala coordinación o lapsos atencionales prolongados, el número de alcances cae.

| user_id | touches_count |
|---------|---------------|
| 1 | 80 |
| 2 | 71 |
| 3 | 61 |
| 4 | 58 |
| 5 | 73 |
| 6 | 70 |
| 7 | 71 |
| 8 | 73 |
| 9 | 60 |
| 10 | 70 |
| 11 | 62 |
| 12 | 26 |
| 13 | 38 |
| 14 | 27 |
| 15 | 26 |
| 16 | 53 |
| 17 | 44 |
| 18 | 35 |
| 19 | 32 |
| 20 | 20 |

Tabla 6: Número de contactos registrados por usuario

Para este IDE, su utilidad resalta porque integra:

- Atención (mantener el estímulo como objetivo),
- Planificación motora (ajustar trayectoria del cursor),
- Corrección online (capacidad de "reenganche").

En neurodegeneración, deterioros en estas funciones (*p.ej brandicinesia, dismetría, fallos ejecutivos*) pueden reflejarse como reducción de alcances y/o mayor irregularidad del patrón de contacto. Cómo se mencionó previamente "este IDE no es estándar en literatura como métrica primaria, pero está fundamentado por coordinación ojo-mano y tareas de tracking manual".

En la Figura 16 se observa la relación entre la edad y el número de contactos registrados durante los 60 segundos de prueba. Con un coeficiente de determinación $R^2 = 0,6263$, existe una relación lineal moderadamente fuerte entre ambas variables dentro del conjunto analizado. El resultado obtenido, implica que, a medida que aumenta la edad, disminuye el número de contactos registrados por prueba, sugiriendo así un claro indicador a tener en cuenta. Además, si en vez efectuar el análisis de la gráfica por tendencia lineal, se hace por tramos, se observa un patrón ya repetido en los IDES anteriores, y es que, a partir de los 35 años, se produce un salto notorio en la disminución del número de contactos. Esto se ve resaltado al que en los usuarios más jóvenes los impactos oscilan entre 62 y 80, sin embargo, al pasar los 35 años el número de impactos pasa a ser de un total de 44 impactos. Este patrón aporta coherencia a la interpretación funcional de este IDE y se complementa con el siguiente IDE.

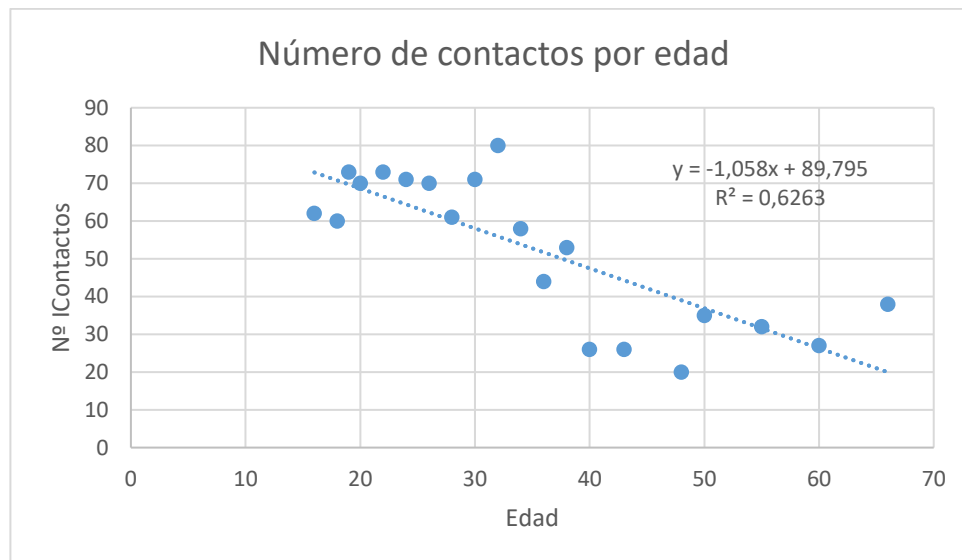


Figura 16: Número de contactos registrados por usuario

4.8. IDE 4: Tiempo efectivo en objetivo

El IDE 4 cuantifica el tiempo real durante el cual el cursor permanece dentro del área del estímulo, aportando un punto de vista temporal al análisis del seguimiento. Para ello se emplean tanto la fracción de tiempo en objetivo, como la duración media de cada episodio de contacto.

La Tabla 7 muestra los valores tenidos en cuenta para el análisis de este IDE.

| user_id | on_target_fraction | mean_contact_duration_s | contact_radius_px |
|---------|--------------------|-------------------------|-------------------|
| 1 | 0,560976 | 0,418048 | 38 |
| 2 | 0,576371 | 0,487657 | 38 |
| 3 | 0,631948 | 0,62057 | 38 |
| 4 | 0,498566 | 0,509637 | 38 |
| 5 | 0,646145 | 0,530502 | 38 |
| 6 | 0,734468 | 0,630869 | 38 |
| 7 | 0,673265 | 0,5691 | 38 |
| 8 | 0,648497 | 0,530302 | 38 |
| 9 | 0,719796 | 0,71421 | 38 |
| 10 | 0,663384 | 0,570839 | 38 |
| 11 | 0,620056 | 0,601133 | 38 |
| 12 | 0,119378 | 0,273977 | 38 |
| 13 | 0,130278 | 0,204227 | 38 |
| 14 | 0,109192 | 0,24697 | 38 |
| 15 | 0,166805 | 0,382507 | 38 |
| 16 | 0,276471 | 0,308754 | 38 |
| 17 | 0,231744 | 0,316652 | 38 |
| 18 | 0,169596 | 0,29289 | 38 |
| 19 | 0,136693 | 0,254993 | 38 |
| 20 | 0,116195 | 0,3461 | 38 |

Tabla 7: Métricas usadas para el análisis del tiempo efectivo en objetivo

En la Figura 17 se muestra la dispersión de la fracción del tiempo en objetivo por edad. En ella se observa que conforme la edad aumenta, el tiempo disminuye, tendencia la cual ya hemos observado en otros IDES. De hecho, ocurre lo mismo, y es que, en usuarios jóvenes (hasta los 35 años), la fracción del tiempo en objetivo se sitúa entre 0,60 y 0,73, lo que implica que el cursor permanece sobre el estímulo durante más del 60% del ensayo. Además, la duración media de los episodios de contacto, los cuales se muestran en la Figura 18 oscila entre 0,50 y 0,71 segundos, reflejando un seguimiento sostenido y estable.

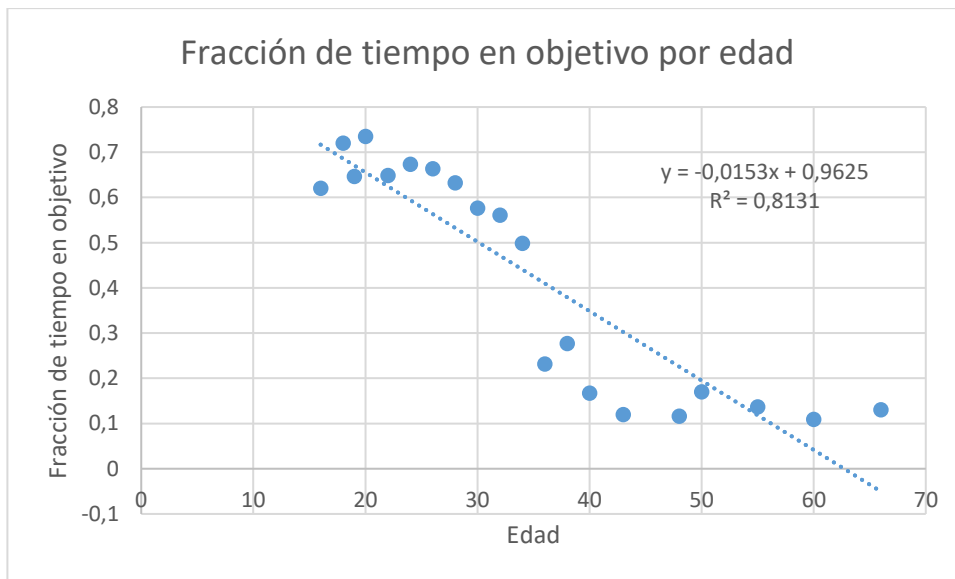


Figura 17: Fracción del tiempo en objetivo de cada usuario ordenados por edad

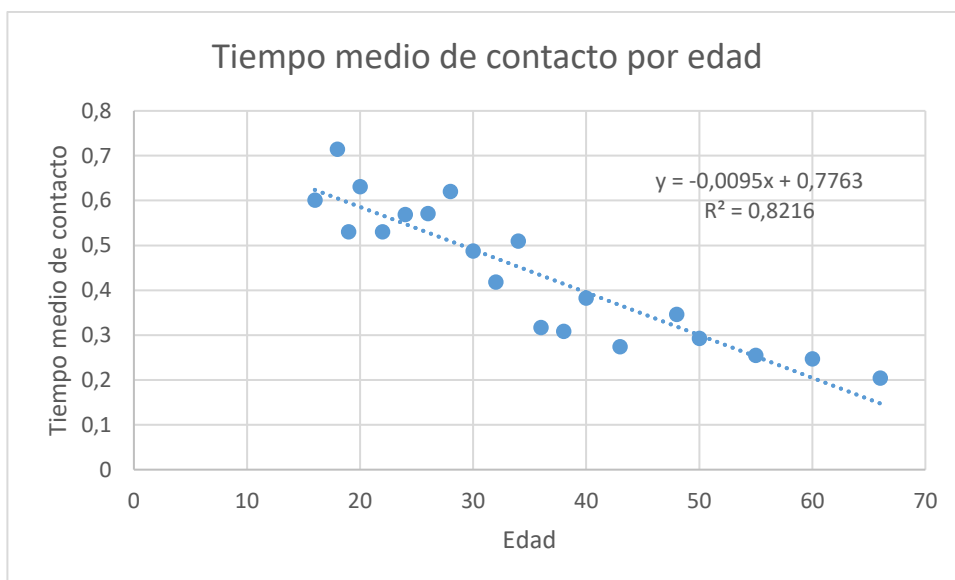


Figura 18: Media del tiempo en objetivo de cada

Por el contrario, los sujetos de mayor edad presentan valores descendientes de forma pronunciada. Se observan fracciones de tiempo en objetivo entre 0,1 y 0,3, junto con duraciones medias de contacto reducidas (entre 0,20s y 0,28s), lo que indica que a medida que aumenta la edad, al igual que disminuye el número de contactos en objetivo, dichos contactos son mucho más breves que aquellos de menor edad.

El análisis de dispersión muestra dicha tendencia, habiendo una correlación de tiempo en objetivo y edad de $R^2 = 0,8131$, y una correlación de duración media de contacto y edad de $R^2 = 0,8216$, evidenciando una relación muy fuerte.

Cabe decir que en comparación con el IDE 3, cuya correlación también es negativa, el IDE 4 presenta una sensibilidad aún mayor, concluyendo de esta manera que, el envejecimiento no solo reduce el

número de alcances exitosos, sino que afecta de forma más marcada a la capacidad de mantener el contacto una vez alcanzado el estímulo, y evidenciando de nuevo ese punto de inflexión visto a partir de los 35 años.

4.9. IDE 5: Latencia en la respuesta verbal al cambio de color.

La latencia en la respuesta verbal cuantifica el tiempo transcurrido desde el cambio de color del estímulo hasta la detección de la identificación mediante voz de ese cambio por parte del usuario. En esta tarea, el usuario debe mantener el seguimiento visomotor del estímulo y, simultáneamente, verbalizar el nuevo color, por lo que el indicador refleja rendimiento verbal bajo condición de doble tarea.

Operacionalmente, para la latencia se tuvieron en cuenta las siguientes métricas: media, mediana, desviación estándar y percentil 95, así como la tasa de éxito dentro de ventana. Además, se monitoriza el número de cambios de color y el número de eventos con latencia válida para asegurar comparabilidad entre registros.

Desde el punto de vista cognitivo, este IDE 5 integra velocidad de procesamiento perceptivo, atención selectiva al cambio, acceso léxico para la denominación del color y planificación/ ejecución del habla. Al tratarse de una situación dual (*persecución + verbalización*), el indicador es especialmente sensible a la disponibilidad de recursos atencionales y al control ejecutivo. Cabe destacar que, en este trabajo, a la hora de separar las latencias por color, si en una prueba el usuario no recita la palabra durante el tiempo que está ese color en pantalla, y se produce un cambio, la latencia para ese color previo no se tiene en cuenta y se considera un error. Esto explica que, en los rebotes del estímulo por las esquinas, dado que son cambios muy rápidos, normalmente se producen dos errores consecutivos.

En la *Tabla 9* se observa una variabilidad notable entre registros, lo que confirma que el sistema discrimina rendimientos verbales distintos. En particular, el uso del percentil 95 resulta clave para detectar episodios de latencia prolongada: pueden aparecer casos con mediana relativamente moderada, pero p95 elevado, lo que sugiere lapsos puntuales (*atencionales/ejecutivos*) más que un enlentecimiento uniforme. Esta lectura es relevante en el contexto del envejecimiento y la neurodegeneración, donde la inestabilidad del rendimiento (*colas largas*) puede ser una señal temprana.

| user_id | n_color_changes | n_within_latency | latency_mean_s | latency_median_s | latency_p95_s | latency_success_rate_within_window |
|---------|-----------------|------------------|----------------|------------------|---------------|------------------------------------|
| 1 | 24 | 21 | 5,791285 | 5,057335 | 15,29913 | 0,291667 |
| 2 | 22 | 19 | 4,961499 | 4,916144 | 10,09596 | 0,272727 |
| 3 | 23 | 20 | 4,556894 | 4,870006 | 11,96377 | 0,304348 |
| 4 | 26 | 23 | 3,27851 | 1,025289 | 15,46362 | 0,576923 |
| 5 | 27 | 25 | 5,374886 | 4,753722 | 13,96101 | 0,259259 |
| 6 | 24 | 21 | 5,830829 | 5,146872 | 12,22695 | 0,25 |
| 7 | 27 | 24 | 5,204835 | 4,699028 | 10,81608 | 0,333333 |
| 8 | 23 | 20 | 5,250788 | 5,193303 | 9,788755 | 0,217391 |
| 9 | 24 | 21 | 4,281776 | 3,692147 | 9,024804 | 0,333333 |
| 10 | 27 | 24 | 10,676 | 9,466397 | 28,60519 | 0,407407 |
| 11 | 26 | 25 | 1,991659 | 0,933583 | 4,920424 | 0,653846 |
| 12 | 30 | 27 | 2,628325 | 1,672316 | 8,177256 | 0,7 |
| 13 | 27 | 26 | 4,055567 | 0,992741 | 20,04795 | 0,703704 |
| 14 | 24 | 23 | 3,288855 | 1,823056 | 9,277712 | 0,541667 |
| 15 | 26 | 26 | 2,625859 | 0,837351 | 11,44208 | 0,807692 |
| 16 | 28 | 26 | 3,930852 | 3,950344 | 8,19314 | 0,428571 |
| 17 | 27 | 27 | 1,20296 | 0,8392 | 2,49587 | 0,962963 |
| 18 | 23 | 23 | 2,815595 | 1,877224 | 7,539885 | 0,565217 |
| 19 | 25 | 22 | 2,44237 | 0,764233 | 12,00398 | 0,72 |
| 20 | 24 | 24 | 2,790645 | 1,945435 | 5,80045 | 0,625 |

Tabla 8: Resumen de la latencia de respuesta verbal ante cambios de color

Con el objetivo de analizar el efecto del estímulo cromático, se evaluaron las latencias por color(rojo/amarillo/verde), como se puede observar en la Tabla 10, donde se muestran las latencias separadas por color.

| user_id | rojo_lat | amarillo_lat | verde_lat |
|---------|----------|--------------|-----------|
| 1 | 5,210042 | 5,999704 | 6,242266 |
| 2 | 3,287979 | 5,087449 | 9,088353 |
| 3 | 5,422549 | 2,932732 | 5,441818 |
| 4 | 2,300904 | 2,442564 | 5,812293 |
| 5 | 5,53475 | 4,759007 | 5,907885 |
| 6 | 4,356849 | 8,037847 | 4,933138 |
| 7 | 3,134219 | 5,571616 | 6,908669 |
| 8 | 4,4237 | 5,137199 | 6,348245 |
| 9 | 3,620546 | 4,596044 | 4,634186 |
| 10 | 8,455097 | 9,917236 | 13,07781 |
| 11 | 1,350855 | 2,481168 | 1,863234 |
| 12 | 2,361157 | 2,465175 | 2,988895 |
| 13 | 2,474613 | 2,87638 | 9,811685 |
| 14 | 2,897957 | 3,869059 | 3,101598 |
| 15 | 1,067507 | 4,310802 | 2,483444 |
| 16 | 4,513514 | 4,234929 | 3,288269 |
| 17 | 1,136031 | 0,779242 | 1,527359 |
| 18 | 2,73506 | 0,39722 | 3,753575 |
| 19 | 3,4374 | 1,042729 | 2,582739 |
| 20 | 4,444335 | 1,766068 | 1,846974 |

Tabla 9: Latencias separadas por color

Para un estudio más preciso, se calcularon las medidas que se observan en la Tabla 11, donde a simple vista se observa que el verde es sistemáticamente el color con mayor media, mayor mediana, mayor p95 y menor tasa de éxito, y donde los colores rojo y amarillo presentan perfiles bastante similares.

| Medias | Rojo | Amarillo | Verde |
|------------------|-------|----------|--------|
| Latencia media | 3.61s | 3.94s | 5.08s |
| Latencia mediana | 2.62s | 2.92s | 3.66s |
| P95 | 8.42s | 9.29s | 11.64s |
| Tasa de éxito | 0.52 | 0.55 | 0.45 |

Tabla 10: Cálculo de medias, para cada medida, de las latencias de respuesta verbal

En la Figura 19, se muestra el gráfico por “boxplots” de la distribución de las latencias en su percentil 95, al igual que se hizo en los IDES 1 y 2, donde se puede observar que la distribución no sigue una progresión lineal ni normal. Estrictamente hablando, se ve que del primer grupo etario (16–25 años) al segundo (26–35 años) se produce un salto enorme en relación con la mediana, para ser exactos, 5s más de latencia. Después, en vez de seguir una tendencia lineal, se produce un punto de inflexión donde, dicha latencia disminuye hasta los 8s, y a partir de aquí, sí que se observa una tendencia ascendente donde el último salto de grupo de etario supone una diferencia de unos 7s con relación a la mediana. Esto rompe la hipótesis de linealidad que llevábamos observando en los anteriores IDES, dado que, no hay una distribución normal, ni un crecimiento monotónico, sino un comportamiento no lineal con inflexión intermedia.

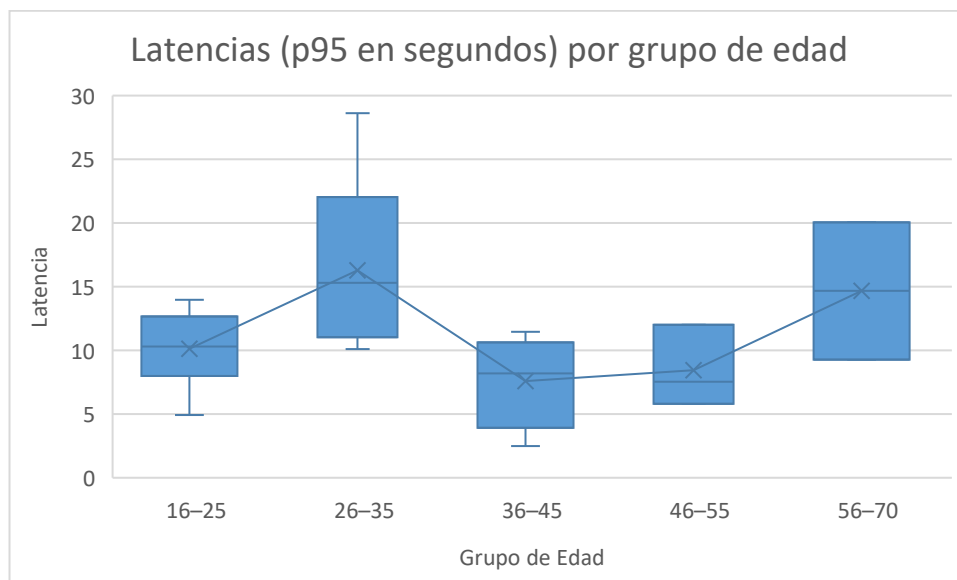


Figura 19: Latencias mostradas por grupo de edad ascendente en su percentil 95

Dado que la distribución general de la latencia por grupo etario no aportaba mucha información al IDE, se realizó el estudio de las latencias separadas por color por edad de usuario en su percentil 95. En la Figura 20, muestra el primer color analizado, en este caso, el color rojo. Aquí, se observa que no hay una pendiente apenas en la recta de regresión (su linealidad no es significativa para el estudio), sino que, los datos mantienen más o menos los mismos valores respecto a la edad siendo el coeficiente de correlación muy bajo ($R^2 = 0,0088$). De esta gráfica, lo único concluyente, es el hecho de que la edad no explica prácticamente nada del percentil 95 rojo, siendo la variabilidad completamente interindividual, lo que sugiere que al menos el color rojo no muestra dependencia con la edad.

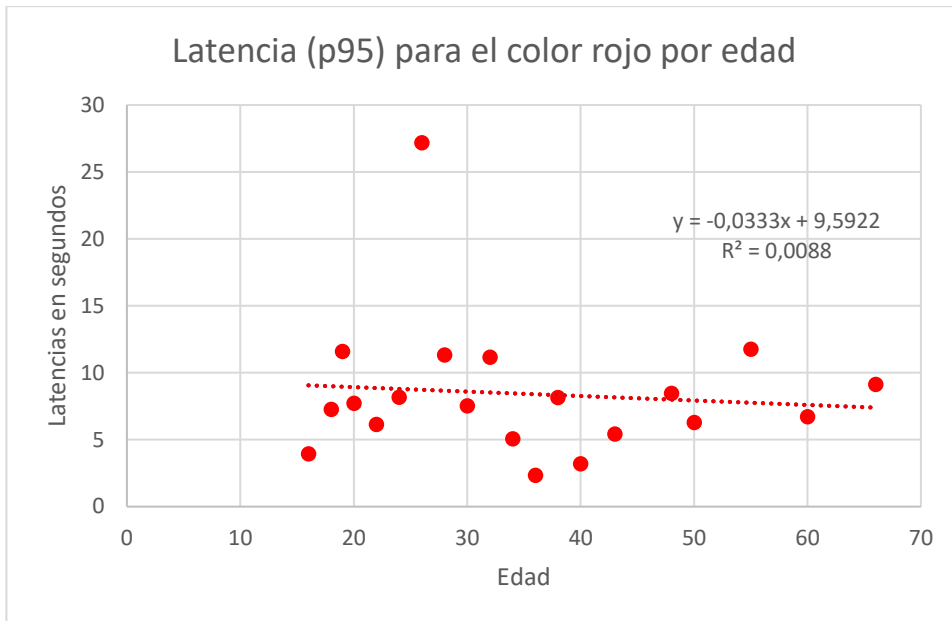


Figura 20: Distribución de la latencia en su percentil 95 para el color rojo

Para el caso mostrado en la Figura 21, pese a que la pendiente de la recta en comparación con el color rojo es algo más pronunciada, su coeficiente de correlación sigue siendo muy bajo ($R^2 = 0,0732$), por lo que, la latencia para el color amarillo sigue sin ser significativa para su estudio en base a la diferencia de edad, observando algunos casos donde la latencia es extrema a edades tempranas.

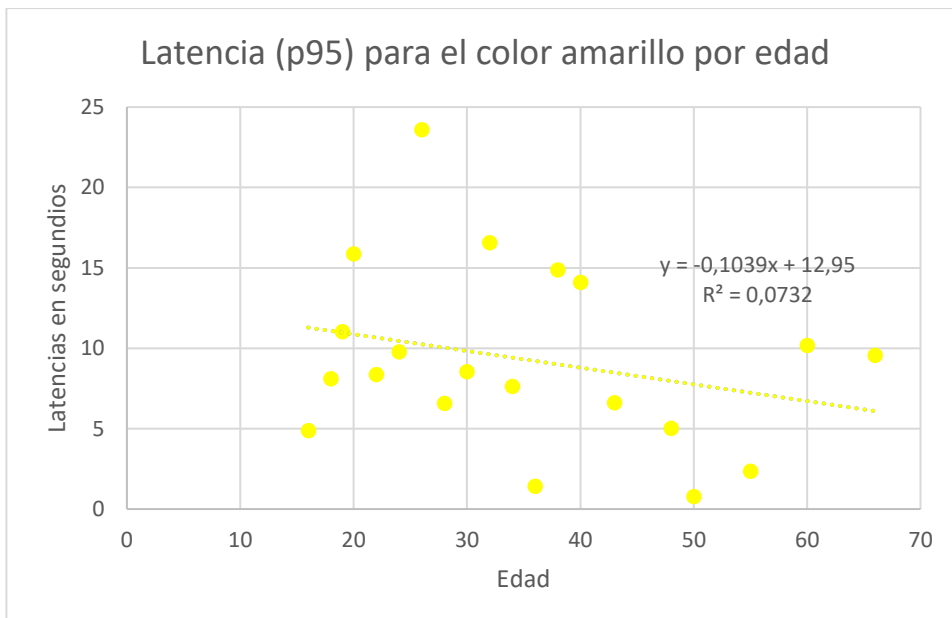


Figura 21: Distribución de la latencia en su percentil 95 para el color amarillo

En la Figura 22, la pendiente vuelve a ser prácticamente plana, por lo que, para el estudio de la latencia para el color verde, la edad vuelve a ser independiente.

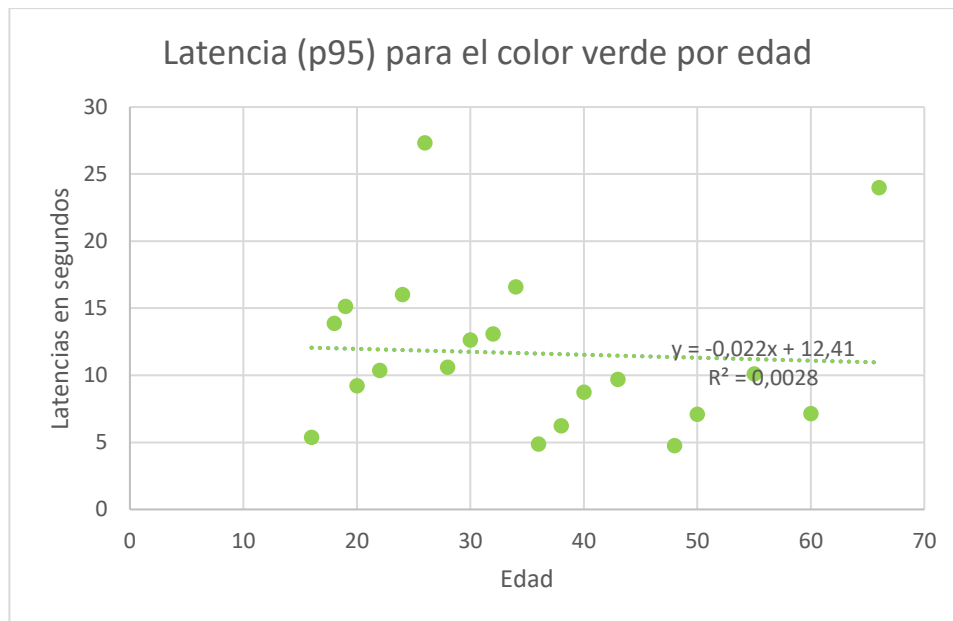


Figura 22: Distribución de la latencia en su percentil 95 para el color verde

En ninguno de los colores la edad explica de manera significativa la latencia extrema. Todos los coeficientes de correlación R^2 son muy bajos, menores que $<0,08$, lo que indica que no hay dependencia lineal entre la edad y el percentil 95 de las latencias, tanto por color como a nivel general, la variabilidad está dominada por factores individuales a lo largo de los diferentes ensayos, y el patrón observado en la gráfica por grupos no se debe a un gradiente lineal, sino a heterogeneidad interna, aunque cabe destacar que para el color rojo se presenta muy poca dispersión a diferencia de los otros dos.

Dado que las latencias en su percentil 95, no muestran mucha información, en la Figura 23, se presenta la distribución de la media y de la mediana de la latencia conforme a la edad, donde la ecuación superior representa la recta de regresión para la media y la ecuación inferior representa la recta de regresión para la mediana, ambas con sus coeficientes de correlación correspondientes. En dicha figura, ambas pendientes son negativas, y los coeficientes de correlación siguen siendo insignificantes, lo que confirma que ni la media ni la mediana presentan relación con la edad, aportando menos información que el estudio previo por percentil 95.

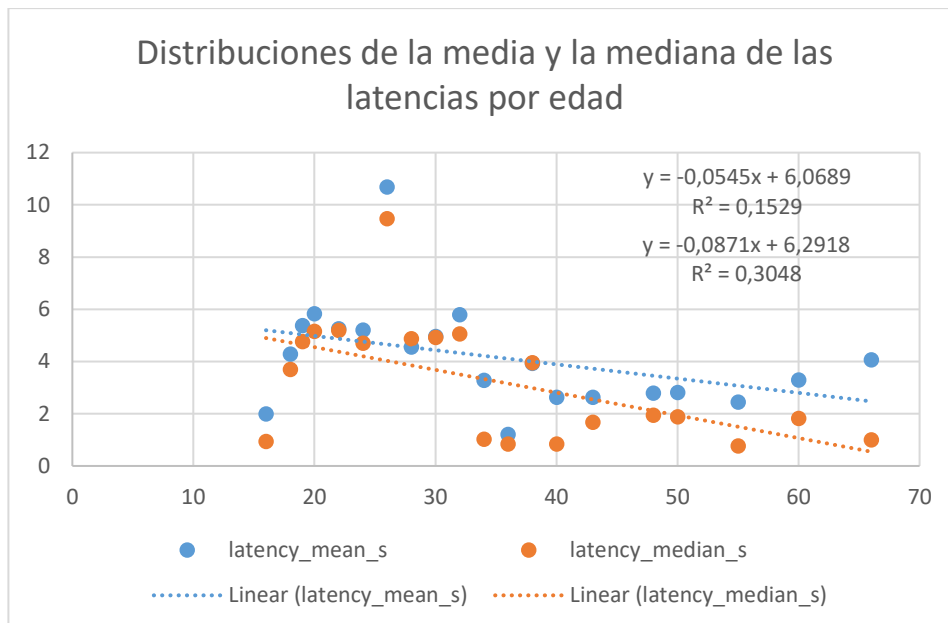


Figura 23: Distribución media y mediana de la latencia por edad

Tras observar la independencia de la edad respecto a la latencia, lo único concluyente es que, a diferencia de otros factores mostrados en los anteriores IDES, la capacidad para reconocer colores y recitarlos en voz alta no se pierde a medida que avanza el tiempo en las personas.

4.10. IDE 6: Precisión verbal (aciertos/errores en denominación)

La precisión verbal cuantifica el grado de concordancia entre el color real mostrado por pantalla en el estímulo y el color reconocido por voz a partir de la transcripción de *Whisper*. A diferencia del IDE 5 (latencia), que mide *cuándo* responde el usuario, el IDE 6 mide si se responde correctamente. Por tanto, ambos indicadores son complementarios, ya que un usuario puede responder rápido, pero equivocarse, o responder lento, pero acertar.

En este estudio se han utilizado dos medidas. La primera es *accuracy_voice_nonzero*, la cual muestra la precisión calculada solo en instantes donde hay detección de voz válida, es decir, hay momentos donde el usuario no está hablando durante su ensayo, por lo que automáticamente se rellenan Os para mostrar estos instantes, por lo que esta en esta medida se eliminan dichos Os. Captura la exactitud cuando el usuario responde. La segunda es *accuracy_all_frames*, la cual muestra la precisión calculada sobre todos los instantes, penalizando silencios u omisiones. Captura tanto exactitud como disponibilidad de respuesta. Esta dualidad permite separar entre errores de denominación y omisiones/no respuesta/no detección.

En la Tabla 12, se muestran estas medidas, y se observa que *accuracy_voice_nonzero* oscila entre 0,47 y 0,83 aproximadamente, y *accuracy_all_frames*, entre 0,46 y 0,83, lo que confirma que el

IDE 6 discrimina ejecuciones con distinta exactitud verbal. Pese a esto, en la mayoría de los casos, ambas métricas muestran valores parecidos, es decir, el descenso por incluir todos los frames es pequeño. En muchos casos el gap es 0, por lo que coinciden o el máximo valor que alcanza es de 0,0285 para el caso del usuario 15, por lo que ambas métricas son casi exactas.

| user_id | accuracy_voice_nonzero | accuracy_all_frames |
|----------------|-------------------------------|----------------------------|
| 1 | 0,688153 | 0,688153 |
| 2 | 0,647008 | 0,626442 |
| 3 | 0,726839 | 0,726839 |
| 4 | 0,732932 | 0,732932 |
| 5 | 0,68901 | 0,68901 |
| 6 | 0,72539 | 0,72539 |
| 7 | 0,665735 | 0,664065 |
| 8 | 0,692891 | 0,692891 |
| 9 | 0,722462 | 0,712422 |
| 10 | 0,546929 | 0,546929 |
| 11 | 0,763842 | 0,763842 |
| 12 | 0,471347 | 0,456691 |
| 13 | 0,804167 | 0,804167 |
| 14 | 0,737326 | 0,737326 |
| 15 | 0,631167 | 0,60266 |
| 16 | 0,829972 | 0,829972 |
| 17 | 0,630744 | 0,624139 |
| 18 | 0,635335 | 0,610879 |
| 19 | 0,737472 | 0,737472 |
| 20 | 0,650582 | 0,650582 |

Tabla 11: Variables empleadas en el cálculo de la precisión verbal

Los resultados del *dataset* (Tabla 12) muestran variabilidad entre registros en ambas medidas. En términos generales, *accuracy_voice_nonzero* se sitúa aproximadamente entre 0,47 y 0,83, lo que confirma que el rendimiento no es plano y que el IDE discrimina ejecuciones con distinta exactitud. Además, la diferencia entre *accuracy_voice_nonzero* y *accuracy_all_frames* es, en la mayoría de los ensayos, relativamente pequeña, lo que sugiere que gran parte de la degradación del rendimiento proviene de errores de identificación más que de omisiones prolongadas. Sin embargo, en algunos

casos se aprecia una brecha mayor entre ambas medidas, indicando que parte de la pérdida de precisión se explica por episodios donde no se produce respuesta válida (*vacilaciones, silencios, o reconocimiento insuficiente*), un patrón relevante porque en contextos clínicos puede corresponder a dificultades de iniciación del habla o a fallos atencionales ante el cambio de estímulo.

La Figura 25, muestra de manera visual, lo ya explicado para la tabla 12, donde se observa que ambas medidas no difieren apenas entre sí.

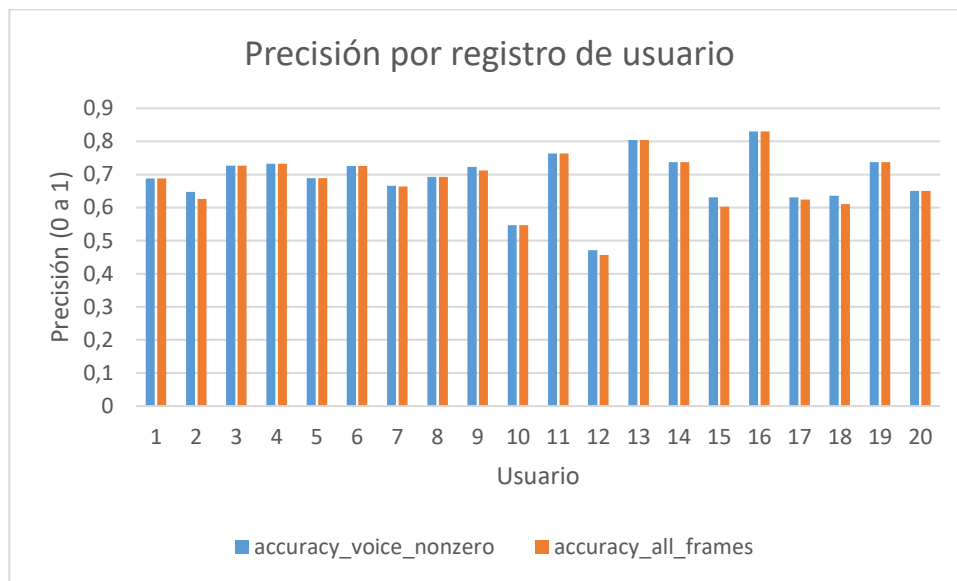


Figura 24: Comparación directa de la precisión por registro de usuario (valores entre 0 y 1)

En la Figura 25, se observa una nube de puntos relativamente estable a lo largo del rango de edades. El ajuste lineal indicado en la propia figura muestra un coeficiente de correlación $R^2 = 0,0056$ lo cual implica que la edad explica un 0,56% de la variabilidad observada en la precisión, por lo que la relación de la edad con este IDE no es significativa a primera vista. Además, la pendiente es prácticamente inexistente por lo que no hay mucho que decir aparte de que apenas hay dispersión en la precisión entre los diferentes sujetos.

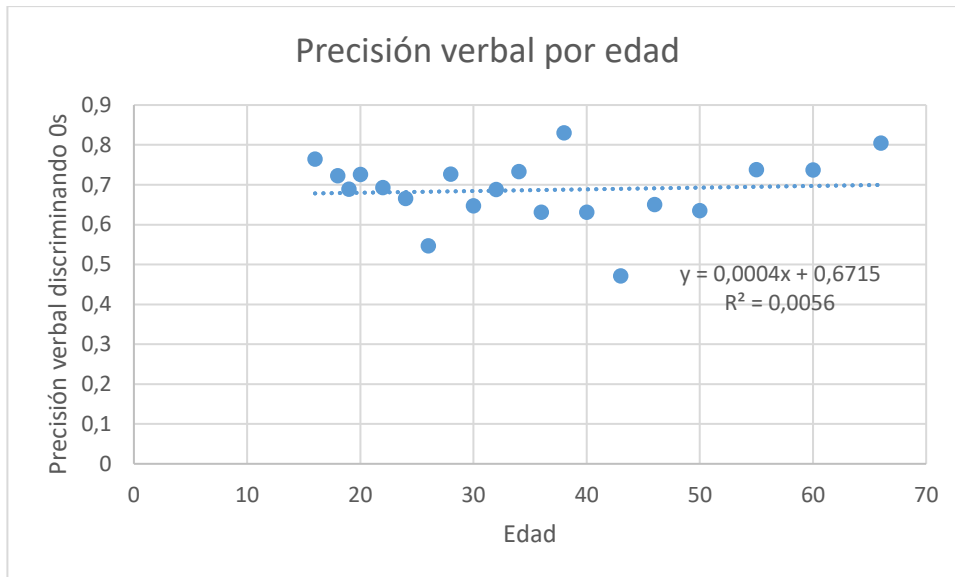


Figura 25: Precisión verbal por edad

En la Figura 26, se muestra la tasa de éxito para el color rojo, donde se aprecia una tendencia, en edades mayores aumenta la probabilidad de responder dentro de ventana, pero la dispersión indica que hay usuarios jóvenes con tasas altas y usuarios mayores con tasas medias/bajas. Por tanto, el envejecimiento sigue sin determinar el desempeño en esta tarea, pero sí se asocia de forma exploratoria con una mejora en consistencia de respuesta en rojo.

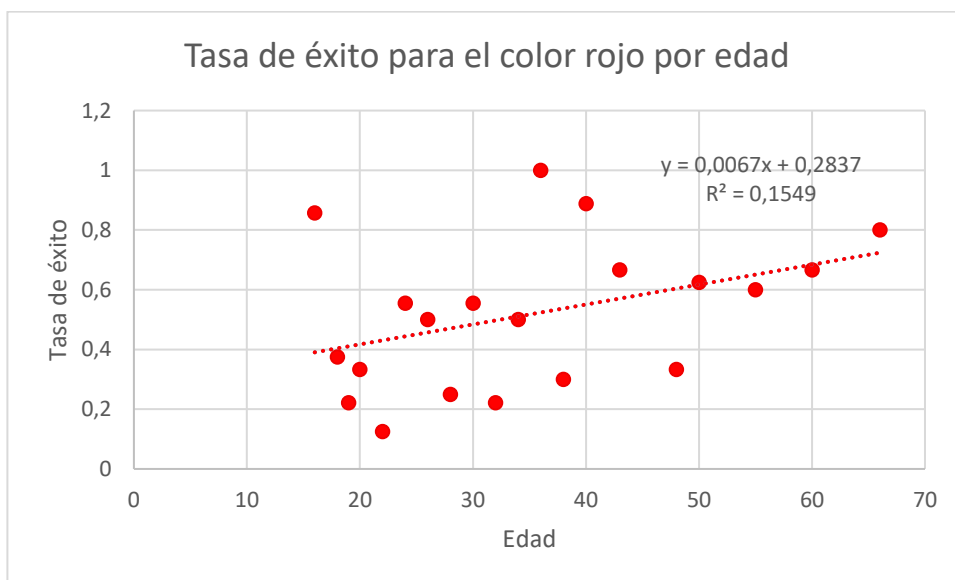


Figura 26: Tasa de éxito para el color rojo

En la Figura 27, se muestra el caso para el color amarillo, este resultado es el más relevante del IDE 6, ya que el amarillo presenta una asociación lineal con la edad bastante superior al resto. En el dataset, conforme a la edad, aumenta de manera clara la tasa de éxito dentro de ventana para el amarillo. Dicho esto, el resultado mostrado, no implica causalidad, sino que puede estar influido por

el tamaño muestral reducido o por casos concretos, pero aun así, es el patrón más consistente para esta métrica, lo que puede indicar que el amarillo es el color que mejor reconocen las personas.

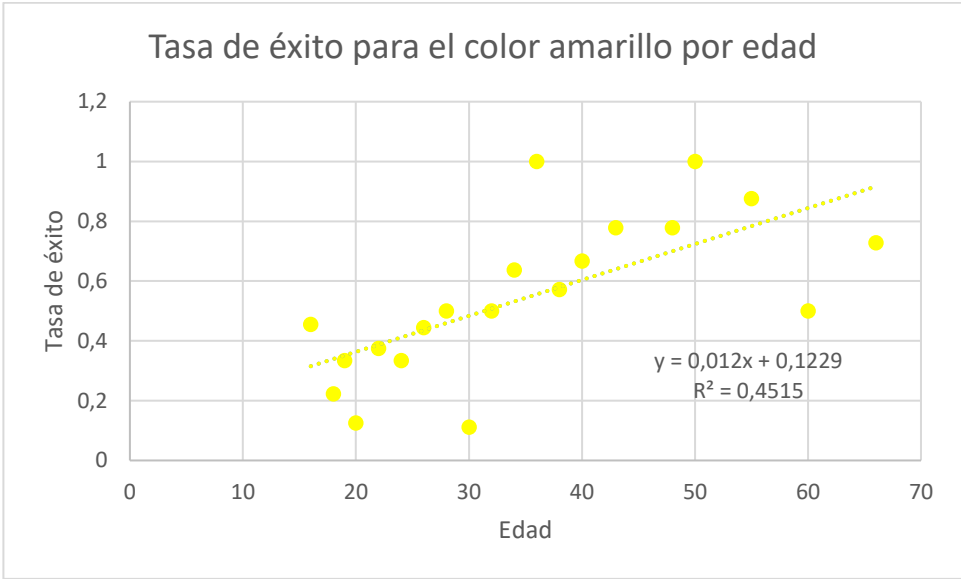


Figura 27: Tasa de éxito para el color amarillo

En el caso de la Figura 28, para el color verde también hay una tendencia positiva, pero su dispersión es amplia. Esto sigue respaldando el hecho de que la edad explica poco la consistencia de respuesta en verde, donde hay perfiles jóvenes con éxito muy bajo y otros con éxito medio-alto, y lo mismo en mayores.

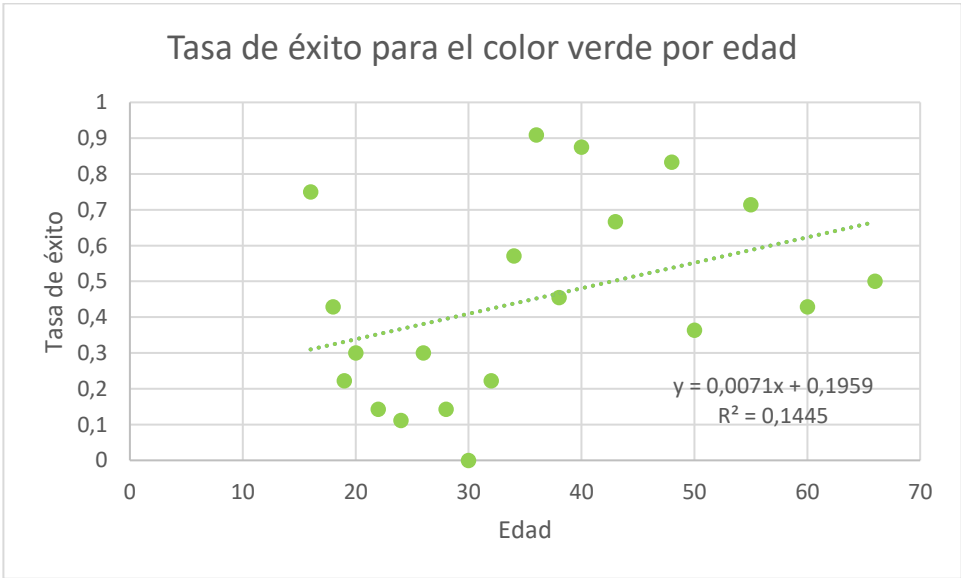


Figura 28: Tasa de éxito para el color verde

En este trabajo, la edad no parece modular la exactitud de denominación en la respuesta verbal ante estímulo, pero sí se asocia (de forma exploratoria), con la consistencia temporal de respuesta dentro

de ventana, sobretodo, para el color amarillo, ya que en los resultados individuales de cada usuario, los cuáles han sido registrados pero no tomados en cuenta para el análisis de este trabajo, no había alteraciones en dicha respuesta a lo largo del ensayo, lo que puede indicar que la fatiga temporal no ha sido suficiente o que realmente el hecho de reconocer un color, no supone un esfuerzo notorio. La combinación de latencia (*IDE 5*) y precisión (*IDE 6*) permite distinguir perfiles funcionales distintos (*p. ej., rápido pero impreciso frente a lento pero preciso*), especialmente relevantes en estudios de envejecimiento y en la investigación de marcadores digitales para enfermedades neurodegenerativas. [15] Para ser exactos, en el *IDE 5*, el verde tiende a presentar latencias más altas y peor rendimiento temporal, mientras que en el *IDE 6* el verde muestra tasas de éxito más bajas con mayor dispersión, por lo que, a nivel funcional, latencia y éxito temporal tienden a comportarse en sentido opuesto: cuando aumentan las latencias, disminuye la probabilidad de respuesta dentro de ventana.

4.11. Cierre de los Resultados Obtenidos

El análisis de los IDES realizado en este trabajo permite concluir que la tarea propuesta de seguimiento de un estímulo dinámico mediante cursor y respuesta verbal simultánea genera métricas capaces de capturar patrones funcionales relevantes. En particular, los resultados muestran dos bloques claramente diferenciados:

- Indicadores asociados al control visomotor y a la interacción efectiva con el objetivo (*IDE1-IDE4*), que sí evidencian tendencias con la edad.
- Indicadores asociados al procesamiento verbal por color (latencia y precisión) cuyo comportamiento se mantiene más estable frente a la edad dentro del conjunto analizado

En primer lugar, los resultados del *IDE1* (error de seguimiento) y el *IDE2* (variabilidad angular) muestran una tendencia consistente, ya que, a medida que aumenta la edad, el desempeño en la tarea de seguimiento se deteriora, reflejándose en un mayor error y en una mayor variabilidad angular en la trayectoria del cursor al perseguir el estímulo. Esto sugiere que los participantes de mayor edad presentan un control menos preciso y estable, con trayectorias más irregulares (mayores desviaciones angulares) al intentar mantener el seguimiento continuo. Además, el análisis conjunto mediante la gráfica de dispersión refuerza esta interpretación, donde, se observa que, a mayor error de seguimiento, mayor variabilidad angular, y que esta relación se acentúa en perfiles de mayor edad, exceptuando un par de casos anómalos que se han dado, donde a un gran error de seguimiento, la variabilidad angular no se ve afectada, pero al ser casos especiales, no aportan información al

conjunto. Por tanto, IDE1 e IDE2 no solo resultan sensibles a cambios asociados al envejecimiento, sino que también muestran coherencia interna entre métricas: ambos describen el mismo fenómeno desde dos perspectivas complementarias (precisión global del seguimiento y estabilidad/consistencia del control).

En segundo lugar, los indicadores de interacción con el objetivo (IDE3 e IDE4) también muestran resultados informativos y, de hecho, aportan un valor añadido clave: describen el rendimiento no solo como "seguir bien" el estímulo, sino como lograr alcanzar el objetivo y sostenerlo. Los datos evidencian que, con el aumento de la edad, tiende a producirse una reducción del número de contactos y, además, una disminución del tiempo efectivo en objetivo. Esta relación es especialmente relevante porque ambos IDES están conectados de forma natural: los usuarios que realizan menos contactos tienden también a realizar contactos más breves, mientras que los usuarios con mayor número de contactos suelen acumular un tiempo efectivo en objetivo superior. En conjunto, IDE3 e IDE4 capturan un patrón funcional robusto, donde hasta aproximadamente los 35 años se observa un desempeño más alto en términos de contacto (más contactos y mayor duración), mientras que a partir de ese rango aparece una tendencia descendente.

Un punto importante del estudio es la presencia de un "gap" o salto alrededor de los 35 años, que aparece de forma convergente en los indicadores visomotores (IDE1–IDE2) y en los indicadores de contacto (IDE3–IDE4). Este hallazgo debe interpretarse con prudencia, ya que, con el tamaño muestral y la naturaleza del estudio, no es posible afirmar una causa específica ni asociarlo directamente a un proceso patológico. Sin embargo, sí puede plantearse como una señal exploratoria potencialmente relevante. Existen varias explicaciones plausibles que deben considerarse en futuras investigaciones: (a) un efecto del muestreo (pocos participantes en ese rango o alta variabilidad individual), (b) diferencias en estrategias de ejecución o familiaridad con el uso del ratón/ordenador, (c) efectos de fatiga/atención durante la tarea, o (d) cambios graduales en la eficiencia de mecanismos como la coordinación visomotora, la estabilidad atencional o el control fino, que podrían manifestarse antes o de forma más acusada en determinados perfiles. En cualquier caso, el valor científico del "gap" en este trabajo reside que el sistema desarrollado en *Python* es capaz de detectar un punto de inflexión como este, y expresarlo en métricas cuantificables, lo que justifica la necesidad de replicarlo en un estudio más amplio, equilibrado por edades y con mayor control experimental.

Por otro lado, los indicadores asociados al análisis por color (latencia verbal y precisión verbal) muestran una variación más limitada respecto a la edad en el conjunto evaluado. Esto sugiere que, al menos en esta tarea concreta, la capacidad de reconocer un color y nombrarlo en voz alta se

mantiene relativamente estable frente al envejecimiento, o bien que la tarea no es lo suficientemente exigente como para revelar diferencias claras con el tamaño muestral actual. Aun así, aparece un resultado específico consistente: el color verde tiende a asociarse con mayor latencia y menor precisión verbal. Este efecto puede estar relacionado con factores perceptuales, atencionales o de discriminación cromática en el contexto de la interfaz, pero nuevamente requiere confirmación con más ensayos, mayor número de participantes y un diseño que controle mejor variables como iluminación, calibración de pantalla y condiciones de grabación. Por tanto, estos IDES verbales no quedan invalidados; más bien, se concluye que su poder discriminativo con la edad es menor en este estudio y que su utilidad probablemente aumentaría con un *dataset* más grande y un protocolo más robusto.

Además, los resultados señalan que los IDES más prometedores para capturar tendencias asociadas a la edad son los cuatro primeros: IDE1 (error de seguimiento), IDE2 (variabilidad angular), IDE3 (número de contactos) e IDE4 (tiempo efectivo en objetivo). Además, se demuestra que existen relaciones coherentes entre pares de indicadores (IDE1–IDE2 y IDE3–IDE4), lo cual es particularmente valioso porque refuerza la interpretación: no se trata de métricas aisladas, sino de un conjunto que describe de forma consistente el desempeño desde dimensiones complementarias (precisión, estabilidad, éxito y persistencia en el objetivo). Esta convergencia sugiere que la tarea propuesta es sensible a cambios funcionales que pueden estar relacionados con envejecimiento y, en un futuro, podría explorarse su capacidad para diferenciar perfiles con alteraciones más marcadas. En conclusión, se efectuó la validación del enfoque técnico: la API desarrollada en *Python* permite integrar un estímulo visual dinámico, adquirir el movimiento del cursor, capturar el audio y sincronización temporal, y obtener de forma automatizada información sobre las métricas utilizadas para el análisis de los ya explicados IDES a partir de registros estructurados. Esto demuestra que la aplicación constituye una plataforma sólida y reproducible para estudios posteriores, ya que facilita:

- Ampliar el tamaño muestral
- Repetir el protocolo en entornos distintos
- Incorporar nuevos indicadores

En consecuencia, la herramienta es válida tanto como demostración experimental, como una base práctica para futuras líneas de investigación orientadas a detectar patrones de desempeño asociados a pérdida funcional, envejecimiento o incluso alteraciones neurocognitivas, siempre planteándolo como evaluación objetiva del rendimiento y no como diagnóstico clínico directo.

5

Conclusiones y líneas futuras

El presente TFG demuestra que una API experimental en *Python* como la desarrollada (control del estímulo, registro del cursor, captura y transcripción de voz, sincronización temporal y exportación estructurada) es útil y válida como infraestructura para generar *datasets* multimodales reproducibles y extraer Indicadores de Desempeño (IDES) comparables entre sujetos en un contexto de evaluación funcional no invasiva y de bajo coste [1][2] [3] [8].

Aunque nuestra contribución principal es el propio diseño de esta API, se han definido un conjunto IDES, y se ha planificado un estudio experimental con sujetos reales, para extraer dichos identificadores. De esta forma, hemos obtenido como primer resultado, una validación técnica de la herramienta, y como segundo resultado, una valoración de la significación de los IDES desde un punto de vista del desempeño de las personas, que en un futuro y con un asesoramiento adecuado, podría ser utilizado por personal clínico especializado.

A partir de los resultados obtenidos tras su análisis, se puede decir que:

- Los IDES visomotores capturan diferencias compatibles con la edad en esta muestra, mostrando un patrón consistente de empeoramiento del control visomotor (error de seguimiento, variabilidad angular, nº de contactos y tiempo efectivo en objetivo) al aumentar la edad [10] [9] [8].

- Los IDES verbales no muestran una dependencia robusta con la edad en el protocolo actual. En los datos obtenidos, la edad no parece discriminar entre sujetos, aunque cómo ya se comentó en el IDE 5, habría que modificar ciertos aspectos de la prueba, para ver si realmente esto se ciñe a lo visto [12] [16].

Por tanto, de manera estricta la plataforma es más sensible y consistente midiendo desempeño visomotor, que midiendo desempeño verbal por edad con el *pipeline* actual. Esto no invalida la dimensión verbal, pero sí indica que requiere un refuerzo metodológico antes de usarse como indicador estable.

5.1. Tendencias observadas

- Deterioro visomotor con la edad. Los indicadores visomotores muestran un patrón coherente con la literatura previa: en tareas de coordinación visomotora y control fino, el rendimiento tiende a degradarse con la edad (menor precisión, menor eficiencia y mayor variabilidad) [9] [10] [5]. La coherencia entre los IDES del 1 al 4 es relevante porque no depende de una sola métrica, sino de un conjunto que describe precisión, estabilidad y contacto funcional de forma complementaria [7].
- Cambio de régimen observado alrededor de 35 años. Aparece un salto notable entre rangos etarios en varias métricas, aunque el distinto número de individuos en cada rango puede ser un factor que afecte a este resultado.
- Aspectos relacionados con la respuesta verbal: la edad no explica de forma estable la latencia ni la precisión en la respuesta a distintos colores. En tareas de denominación/latencia verbal se han descrito cambios asociados a alteraciones neurocognitivas y motoras del habla [11]. Aun así, en esta implementación concreta, los IDES verbales se ven condicionados por la precisión del reconocimiento automático y por el método de alineación temporal, lo que limita la sensibilidad de los IDES 5 y 6 para detectar tendencias por edad. En consecuencia, cualquier patrón por color debe interpretarse como secundario y potencialmente dependiente de factores perceptivos o experimentales [15].

5.2. Limitaciones del estudio

- Tamaño muestral y equilibrio por grupos. El análisis por edad es exploratorio, por lo que el por lo que el desequilibrio del número de sujetos por grupo limita la potencia del estudio estadístico.

- Estrictamente hablando, el cursor no mide el seguimiento ocular sino el desempeño funcional mano-objetivo, por lo que implica que los IDES solo reflejan coordinación visomotora funcional y no oculomotricidad pura [6].
- Limitaciones en la respuesta verbal. La latencia y la precisión dependen del reconocimiento de voz y de las trayectorias marcadas por el estímulo, que tienen un carácter aleatorio. Sin segmentación robusta la métrica podría incorporar ruido no biológico (errores de transcripción, fallos en la CPU, etc.).

5.3. Mejoras e implementaciones futuras

Como continuidad natural del trabajo, se plantean varias líneas de mejora con alto impacto:

1. Refuerzo estadístico del análisis por edad: ampliar el tamaño muestral, equilibrar rangos etarios y aplicar modelos estadísticos que permitan cuantificar con mayor rigor la relación edad-IDES (correlaciones, regresiones y control de covariables).
2. Extensión a líneas de investigación clínicas: aplicar el mismo protocolo a grupos con diagnóstico (p. ej., deterioro cognitivo leve, Parkinson u otras patologías), comparando distribuciones con la población sana y evaluando sensibilidad de los IDES. [14]
3. Instrumentación oculomotora directa: incorporar seguimiento visual para validar el proxy manual y descomponer qué parte de la variabilidad proviene de control ocular frente a control manual. [4] [6]
4. Mejorar la evaluación de la respuesta verbal: integrar mecanismos de detección de actividad de voz (VAD), estrategias de alineación temporal más precisas, evaluación sistemática de errores de reconocimiento automático de voz para reducir ruido y ajustar el protocolo para evitar que los cambios de color generados por rebotes en las esquinas ocurran con una cadencia excesivamente alta, ya que pueden registrarse como errores o producir latencias no representativas al no asociarse correctamente la respuesta verbal al evento de cambio los IDE 5 y 6. [13]
5. Nuevos IDES orientados a fatiga y dinámica intraensayo: incorporar métricas que capturen degradación temporal dentro del ensayo (fatiga), reenganche tras cambios de estímulo o estabilidad por segmentos, reforzando la sensibilidad a cambios sutiles.
6. Incluir varios niveles de dificultad y comparar persecuciones normales con persecuciones con voz para obtener análisis más informativo que la latencia absoluta o incluso validaciones cruzadas con electrooculografías para saber cuánto del IDE proviene del control ocular y cuánto del manual [6].

En conclusión, con el protocolo y datos actuales, la API es válida como plataforma accesible para evaluación funcional del desempeño visomotor por edad, ya que los IDES 1-4 muestran tendencias consistentes y complementarias. La rama verbal (IDES 5 y 6) no es todavía un bloque estable para detectar tendencias por edad; antes de usarla como indicador sensible, requiere mejoras técnicas en detección de voz, alineación temporal y estandarización más robusta de las condiciones. Por último, el sistema constituye una base sólida, pero el salto a inferencias más fuertes exige mayor tamaño muestral control de covariables y un diseño con niveles distintos de dificultad [8] [10] [21].

6

REFERENCIAS

- [1] G. Joya Caparrós, R. Ávila Ávila, R. García Bermúdez y J. Medrano Montero, «Red universitaria para la investigación y docencia en técnicas de captación y procesamiento de señales electrooculográficas y de fijación de la mirada orientadas a la detección de enfermedades neurodegenerativas y a la monitorización del aprendizaje,» de *Universidades y agenda 2030: La cooperación universidad–empresa–administración para una sociedad comprometida con los ODS*, Málaga, Editorial Comares, 2023, pp. 347-358.
- [2] World Health Organization, «Dementia,» World Health Organization, 31 March 2025. [En línea]. Available: <https://www.who.int/news-room/fact-sheets/detail/dementia>. [Último acceso: 19 January 2026].
- [3] World Health Organization, «Parkinson disease,» World Health Organization, 9 August 2023. [En línea]. Available: <https://www.who.int/news-room/fact-sheets/detail/parkinson-disease>. [Último acceso: 19 January 2026].
- [4] D. C. Niehorster, W. W. F. Siu y L. Li, «Manual tracking enhances smooth pursuit eye movements,» *Journal of Vision*, vol. 15, nº 15, p. 11, 2015.
- [5] A. Coudiere y F. R. Danion, «Eye-hand coordination all the way: from discrete to continuous hand movements,» *Journal of Neurophysiology*, vol. 131, nº 4, pp. 652-667, 2024.

- [6] R. A. Becerra-García, «Technology for processing saccadic electrooculograms in people with spinocerebellar ataxia type 2,» Universidad de Málaga, Málaga, 2022.
- [7] J.-R. Rizzo, M. Beheshti, T. Naeimi, F. Feiz, G. Fatterpekar, L. J. Balcer, S. L. Galetta, A. G. Shaikh, J. C. Rucker y T. E. Hudson, «The complexity of eye-hand coordination: a perspective on cortico-cerebellar cooperation,» *Cerebellum & Ataxias*, vol. 7, nº 1, p. 14, 2020.
- [8] M. W. Smith, J. Sharit y S. J. Czaja, «Aging, motor control, and the performance of computer mouse tasks,» *Human Factors*, vol. 41, nº 3, pp. 389-396, 1999.
- [9] K. M. Hawkins y L. E. Sergio, «Visuomotor impairments in older adults at increased Alzheimer's disease risk,» *Journal of Alzheimer's Disease*, vol. 42, nº 2, pp. 607-621, 2014.
- [10] S. Hocherman, R. Moont y M. Schwartz, «Recruitment of attentional resources during visuomotor tracking: effects of Parkinson's disease and age,» *Cognitive Brain Research*, vol. 21, nº 1, pp. 77-86, 2004.
- [11] E. Schalling y L. Hartelius, «Speech in spinocerebellar ataxia,» *Brain and Language*, vol. 127, nº 3, pp. 317-322, 2013.
- [12] D. Kempler y M. Goral, «Language and dementia: Neuropsychological aspects,» *Annual Review of Applied Linguistics*, vol. 28, pp. 73-90, 2008.
- [13] C. Roncero, J. Nikelski, S. Probst, A. Fernandez, A. Thiel y H. Chertkow, «The semantic storage loss score: An Algorithm for measuring an individual's level of semantic storage loss due to temporal lobe damage in neurodegenerative disease,» *PLoS ONE*, vol. 15, nº 8, p. e0235810, 2020.
- [14] A. M. G. Meade, L. M. Jackson, R. Gavrilova, J. Stierwalt, J. M. Martinez-Thompson, J. R. Duffy, H. Clark, M. M. Machulda, J. L. Whitwell, K. A. Josephs, R. L. Utianski y H. Botha, «Progressive Apraxia of Speech as a Manifestation of Spinocerebellar Ataxia 2: Case Report,» *Neurology: Genetics*, vol. 10, nº 6, p. e200202, 2024.
- [15] F. J. Moreno-Martínez y I. C. Rodríguez-Rojo, «On Colour, Category Effects, and Alzheimer's Disease: A Critical Review of Studies and Further Longitudinal Evidence,» *Behavioural Neurology*, nº 1, p. 960725, 2015.
- [16] M. C. Silveri, «Contribution of the cerebellum and the basal ganglia to language production: Speech, word fluency, and sentence construction—Evidence from pathology,» *The Cerebellum*, vol. 20, nº 2, pp. 282-294, 2021.

- [17] F. R. Danion, J. Mathew, N. Gouirand y E. Brenner, «More precise tracking of horizontal than vertical target motion with both the eyes and hand,» *Cortex*, vol. 134, pp. 30-42, 2021.
- [18] R. A. Andersen y H. Cui, «Intention, action planning, and decision making in parietal-frontal circuits,» *Neuron*, vol. 63, nº 5, pp. 568-583, 2009.
- [19] K. Dalamagkas, M. Tsintou, Y. Rathi, L. J. O'Donnell, O. Pasternak, X. Gong, A. Zhu, P. Savadjiev, G. M. Papadimitriou, M. Kubicki, E. H. Yeterian y N. Makris, «Individual variations of the human corticospinal tract and its hand-related motor fibers using diffusion MRI tractography,» *Brain Imaging and Behavior*, vol. 14, nº 3, pp. 696-714, 2020.
- [20] E. Tzvi, S. Loens y O. Donchin, «The Role of the Cerebellum in Visuomotor Adaptation,» *The Cerebellum*, vol. 21, nº 2, pp. 306-313, 2022.
- [21] S. Kemper, L. Hoffman, R. Schmalzried, R. Herman y D. Kieweg, «Tracking talking: dual task costs of planning and producing speech for young versus older adults,» *Aging, Neuropsychology, and Cognition*, vol. 18, nº 3, pp. 257-279, 2011.
- [22] Occupational Safety and Health Administration (OSHA), «eTools: Computer workstations—Monitors,» Occupational Safety and Health Administration (OSHA), U.S. Department of Labor, [En línea]. Available: <https://www.osha.gov/etools/computer-workstations/components/monitors>. [Último acceso: 19 January 2026].
- [23] Mayo Clinic, «Office ergonomics: Your how-to guide,» Mayo Clinic, 25 May 2023. [En línea]. Available: <https://www.mayoclinic.org/healthy-lifestyle/adult-health/in-depth/office-ergonomics/art-20046169>. [Último acceso: 19 January 2026].

Apéndice A. Manual de Instalación y Uso de *CHROMATRACKER*

A.1. Descripción general del sistema

ChromaTracker es una aplicación experimental multimodal desarrollada en *Python* que ejecuta una tarea de seguimiento visomotor y respuesta verbal simultánea. Durante un ensayo (*duración fija: 60 segundos*) se presenta un estímulo visual (*bola en movimiento*) que cambia entre tres colores (*rojo, amarillo y verde*). El usuario debe:

- Seguir el estímulo con el cursor (ratón), manteniéndose lo más próximo posible al objetivo.
- Nombrar en voz alta el color del estímulo cuando se produce el cambio.

La aplicación registra y sincroniza de forma automática:

- Trayectoria del estímulo (posición X–Y).
- Trayectoria del cursor (posición X–Y).
- Cambios de color del estímulo.
- Respuesta verbal del usuario (reconocida mediante transcripción automática).

Como salida final, cada ensayo genera un fichero estructurado en formato *Parquet* (*final_XXX.parquet*), utilizado posteriormente para el cálculo de los *IDES* descritos en la memoria.

A.2. Estructura del repositorio

En la raíz del repositorio se encuentran los siguientes elementos principales:

- *main.py*: punto de entrada de la aplicación.
- *eogtalk/gui/logic.py*: bucle principal del experimento, control temporal y guardado de datos.

- `eogtalk/gui/ball.py`: comportamiento del estímulo (movimiento, rebotes, pausas y cambios de color).
- `eogtalk/gui/cursor.py`: registro de posiciones del cursor durante el ensayo.
- `eogtalk/gui/whisper.py`: grabación de audio y extracción de respuestas verbales (colores).
- `eogtalk/gui/whisper_loader.py`: carga del modelo de reconocimiento de voz.
- `parquetFinales/`: carpeta de salida de los *Parquet* finales por ensayo.
- `notebooks/`: cuadernos *Jupyter* para cálculo de IDES y exportación a Excel.
- `requirements.txt` y `makefile`: dependencias y comandos de ejecución.

A.3. Requisitos del sistema

i) Requisitos mínimos

- *Python* 3.11 (recomendado por compatibilidad con dependencias).
- Micrófono funcional configurado en el sistema.
- Sistema operativo: Windows / Linux / macOS.

ii) Requisitos recomendados

- Entorno virtual (*venv*) para aislar dependencias.
- *FFmpeg* instalado (útil en algunos entornos para procesamiento de audio).
- En Linux: librerías *PortAudio* si se usa *sounddevice*.

A.4. Instalación

La instalación puede realizarse mediante dos alternativas: (i) instalación con *uv* (si se utiliza este gestor en el entorno del proyecto), o (ii) instalación estándar con *venv* + *pip*.

i) Opción 1: Instalación con *uv*

1. Situarse en la raíz del repositorio.
2. Ejecutar:
 - ***make install***
3. Para ejecutar la aplicación:
 - ***make run***

ii) Opción 2: Instalación estándar con *venv* + *pip*

1. Crear entorno virtual en la raíz del repositorio:
 - `python -m venv venv`
2. Activar el entorno:

- **Windows (PowerShell):**
 - `venv\Scripts\Activate.ps1`
 - **Windows (CMD):**
 - `venv\Scripts\activate.bat`
 - **Linux/macOS:**
 - `source venv/bin/activate`
3. Instalar dependencias:
 - `pip install -r requirements.txt`
 4. Ejecutar la aplicación:
 - `python main.py`

A.5. Ejecución del experimento

1. Ejecutar la aplicación (`python main.py` o `make run`).
2. La interfaz se inicia en pantalla completa y muestra instrucciones básicas.
3. Iniciar el ensayo (según configuración de la interfaz: tecla o interacción de inicio).
4. Tras una cuenta atrás ("3-2-1"), se inicia la tarea (60 s).
5. Durante el ensayo:
 - **seguir el estímulo con el cursor.**
 - **nombrar en voz alta el color del estímulo (rojo / amarillo / verde).**
6. Para finalizar antes de tiempo:
 - **pulsar ESC.**

Al finalizar el ensayo, la aplicación ejecuta el procesamiento necesario y guarda los datos del registro.

A.6. Archivos de salida y formato de datos

Tras cada ensayo, la aplicación genera un *Parquet* final en la carpeta:

- `parquetFinales/final_001.parquet`
- `parquetFinales/final_002.parquet`
- ...

Estos ficheros contienen las variables necesarias para el cálculo de IDEs. De forma general, el *Parquet* final incluye información temporal y columnas asociadas a:

- Posición del estímulo (X-Y).
- Posición del cursor (X-Y).
- Codificación del color del estímulo.

- Codificación del color reconocido en la respuesta verbal (si aplica).

Además de la salida final, durante la ejecución pueden generarse ficheros temporales (posiciones del estímulo, cursor y cambios de color) y fragmentos de audio, que se utilizan durante el procesado y no forman parte del conjunto final entregado para análisis.

A.7. Cálculo de IDES con *Jupyter Notebook* y exportación a *Excel*

El repositorio incluye notebooks dentro de la carpeta *notebooks/* cuyo objetivo es:

1. Detectar automáticamente la carpeta *parquetFinales/*.
2. Seleccionar un fichero *final_XXX.parquet*.
3. Calcular los IDES definidos en la memoria.
4. Exportar un Excel con:
 - tablas de métricas por ensayo,
 - eventos de color y respuesta verbal,
 - y visualizaciones (gráficas) utilizadas en el análisis.

i) Apertura del entorno *Jupyter*

Se puede abrir *Jupyter* desde la raíz del repositorio con "*jupyter notebook --notebook-dir .*", o, si se dispone del comando configurado "*make nb*".

ii) Flujo general dentro del notebook

El notebook sigue un flujo típico:

- Detección del directorio raíz y listado de *Parquet* disponibles.
- Selección del ensayo (índice o nombre del fichero).
- Cálculo de métricas visomotoras y verbales.
- Exportación de resultados a un *Excel* por ensayo (ubicación definida en el notebook).

A.8. Solución de problemas frecuentes

i) Error: "No se encuentra la carpeta *parquetFinales*"

- Asegurarse de ejecutar el notebook dentro del repositorio.
- Comprobar que la carpeta *parquetFinales/* existe en la raíz y contiene ficheros *final_*.parquet*.

ii) La respuesta verbal no se registra o aparece vacía

- Comprobar que el micrófono funciona y está seleccionado como entrada principal del sistema.
- Revisar permisos de micrófono (especialmente en Windows).
- Verificar el entorno acústico (ruido de fondo puede afectar a la transcripción).

iii) Errores con *sounddevice* / *PortAudio*

- En Linux, puede ser necesario instalar librerías *PortAudio* (por ejemplo, *portaudio19-dev* en distribuciones Debian/Ubuntu).
- Revisar configuración de audio del sistema y permisos.

iv) Problemas con pantalla completa

- La aplicación se ejecuta en pantalla completa por diseño experimental.
- Si se requiere modo ventana para pruebas técnicas, es necesario modificar la configuración de *Pygame* en el módulo de interfaz (ajuste opcional, no necesario para reproducir el flujo del TFG).

A.9. Procedimiento recomendado para reproducir el flujo del TFG

Para replicar el proceso completo tal como se presenta en la memoria:

1. Ejecutar uno o varios ensayos con la aplicación.
2. Verificar que se generan ficheros *final_XXX.parquet* en *parquetFinales/*.
3. Abrir el notebook de análisis y seleccionar el *Parquet* correspondiente.
4. Calcular los IDES.
5. Exportar los resultados a *Excel*.
6. Consolidar los resultados globales a partir de los *excels* generados.



UNIVERSIDAD
DE MÁLAGA

| uma.es

E.T.S. DE INGENIERÍA INFORMÁTICA

E.T.S de Ingeniería Informática
Bulevar Louis Pasteur, 35
Campus de Teatinos
29071 Málaga