

## Sistema de Apoyo Experimental para el Reconocimiento Visual y el Nombrado de Objetos

W.R. Rodríguez Dueñas<sup>1</sup>, H. Marín Manrique<sup>2</sup>, B.J. Hernández Jaramillo<sup>1</sup>, M. Torres Narvaez<sup>1</sup>, E. Galindo Rojas<sup>1</sup>

<sup>1</sup>Escuela de Medicina y Ciencias de la Salud, Universidad del Rosario, Bogotá, Colombia.

<sup>2</sup>Facultad de Ciencias Sociales y Humanas, Universidad de Zaragoza, Campus Universitario de Teruel, España.

### RESUMEN

En la práctica clínica relacionada con la rehabilitación cognitiva, perceptual y del lenguaje, es habitual emplear como estímulos representaciones de objetos en dos dimensiones como dibujos o fotografías para que sean identificados por los pacientes. Ello supone una limitación, puesto que la presentación de objetos en tres dimensiones, dada su mayor riqueza visual, facilita su identificación y podría apoyar procesos de rehabilitación. **Objetivo.** Diseñar una configuración experimental para conocer el grado en que las modalidades de presentación de estímulos afectan el reconocimiento visual de objetos. **Materiales y Métodos.** Se diseñó un sistema de bajo costo para mostrar estímulos en tres modalidades: imágenes, videos y objetos reales. Fueron seleccionados 80 objetos, controlando las variables de manipulabilidad, edad de adquisición, familiaridad, complejidad visual y frecuencia léxica. Se realizó una prueba piloto utilizando el sistema con cinco personas sanas y se midieron los tiempos de reacción en la nominación de los objetos. **Resultados.** Las diferencias en los tiempos de reacción entre las modalidades observadas apoyan la influencia de la modalidad de presentación en el reconocimiento de objetos. **Conclusiones.** Las propiedades de volumen de los objetos permitirían la exploración de características que las fotografías no tendrían; razón por la cual, los tiempos de reacción para los objetos reales y los videos fueron menores. Son valiosos los aportes que ofrecen una alternativa de bajo costo para la visualización y el reconocimiento de objetos.

**Palabras clave:** rehabilitación, reconocimiento de objetos, agnosia visual, diseño en ingeniería.

### Correspondencia:

William R. Rodríguez  
Escuela de Medicina y Ciencias de la Salud, Universidad del Rosario.  
Cra 24 63C 74 Bogotá, Colombia.  
Correo electrónico: [william.rodriguez@urosario.edu.co](mailto:william.rodriguez@urosario.edu.co)

### Fecha de recepción:

14 de octubre del 2015

### Fecha de aceptación:

1 de marzo del 2016

## ABSTRACT

A wide variety of stimuli or objects such as drawings or photographs are used in cognitive perceptual and language rehabilitation. However, the presentation of objects in three dimensions, due to the visual richness, facilitates the identification and could support the rehabilitation process. **Objective.** Design an experimental setup to determine the extent to which the modality of presentation of objects affects the visual recognition of objects. **Materials and methods.** A low-cost experimental system was designed in order to show the stimulus in three ways: images, videos and real objects. Eighty objects were selected and variables as manipulability, age of acquisition, familiarity, visual complexity, agreement on the label and lexical frequency were controlled. A pilot study was conducted using the system in five healthy participants and the reaction time to identify objects was measured. **Results.** The findings reveal that the differences in reaction times between the modalities of presentation support the influence of the mode of presentation in object recognition. **Conclusions.** The volume object properties could allow the exploration of characteristics that the photographs would not have. This may be why the reaction time to identify real objects or videos was lower. Low-cost alternatives for the visualization and object recognition are valuable in cognitive perceptual and language rehabilitation.

**Keywords:** rehabilitation, object recognition, agnosia, engineering design.

## INTRODUCCIÓN

La nominación de objetos es una habilidad fundamental que usamos las personas durante la comunicación verbal y consiste en reconocer un objeto presentado visualmente y mencionar su nombre en voz alta. De allí que los objetos presentados visualmente sean comúnmente usados como estímulos tanto en la investigación de funciones mentales superiores (ej., percepción, memoria y lenguaje) como en la práctica terapéutica. Sin embargo, los estímulos pictóricos empleados varían ampliamente en dimensiones como: 1) la concordancia del nombre (de acuerdo en la forma como los objetos son llamados por diferentes personas y culturas); 2) la edad de adquisición de la etiqueta o nombre del objeto; 3) la complejidad visual; 4) la categoría semántica del objeto y 5) la frecuencia de uso y su familiaridad, entre otros. Es decir, todo lo cual puede afectar el éxito y el tiempo de nombrado de los estímulos visuales. En consecuencia, ha habido en los últimos años

un esfuerzo por normalizar este tipo de variables en bases de datos que puedan ser usadas con fines investigativos y clínicos [1, 2, 3].

En la investigación sobre percepción en sujetos atípicos, como ocurre en el caso de los pacientes agnósicos, el reconocimiento de objetos se ve influenciado por el modo de presentación de los mismos. Los pacientes agnósicos tienen un mayor éxito para reconocer objetos reales en comparación con fotografías o dibujos [4, 5], esto seguramente debido a que los objetos reales proporcionan información tridimensional que resulta clave para ayudar en su identificación [6, 7].

En el reconocimiento de objetos se asocia la entrada visual con una categoría semántica; de allí el interés por estudiar la relación entre percepción y acción, en especial de aquellos objetos accionables versus los no accionables o que requieren de un agarre que no evoca un uso específico. Wolk y colaboradores [8] demostraron que la manipulabilidad de los objetos predecía la probabilidad de reconocimiento de los mismos por parte

de sujetos agnósicos, en el sentido que, a mayor manipulabilidad mayor probabilidad de reconocimiento. Sin embargo, definir el grado de manipulabilidad de un objeto no es tan sencillo. Mientras que para Magnié y colaboradores [9] la manipulabilidad de un objeto viene determinada por el grado en que se puede imitar su uso (pantomima) a través de gestos; otros autores proponen que lo importante en relación a la manipulabilidad es saber hasta qué punto la forma del objeto indica su modo de uso [8]. En este sentido vale la pena recordar que las representaciones motoras asociadas al uso de un objeto podrían activarse ante la visión de dicho objeto y ayudar en su reconocimiento y nombrado. Por ello, la prueba piloto del diseño aquí presentado permite comparar el efecto que para el éxito del nombrado (tasas de aciertos y tiempo de reacción) tiene la modalidad de presentación de los estímulos (imágenes, videos y objetos reales).

En el área de la rehabilitación algunos pacientes, en especial aquellos sobrevivientes de enfermedad cerebro vascular, presentan alteraciones en su capacidad para identificar objetos. Esta situación puede reducir el grado de independencia funcional afectando su relación con el entorno, cuyas exigencias implican el reconocimiento y nombrado de objetos. Los terapeutas disponen de recursos para la rehabilitación que se apoyan mayoritariamente en la presentación de láminas con dibujos o fotografías en dos dimensiones de los objetos. A priori esto parece una limitación, puesto que si la identificación de los objetos se ve facilitada por la riqueza de detalles visuales, resultaría más conveniente presentar los objetos reales y no una representación gráfica de los mismos. Sin embargo, teniendo en cuenta la cantidad ingente de estímulos (objetos) que se necesitan en un proceso de rehabilitación clínica y los problemas derivados de su gestión y almacenamiento, es entendible que se sigan empleando dibujos o fotografías.

Una tendencia de investigación es el

escaneo de objetos reales por medio de tecnologías digitales para su posterior reconstrucción virtual. Como ejemplo de ello se encuentra el escaneo con láser de objetos reales y la utilización de avanzados algoritmos de reconstrucción para obtener al final del proceso, objetos virtuales para diversas aplicaciones [10, 11, 12]. También existen bases de datos libres en internet como la OB3D, una herramienta para la investigación experimental en ciencias cognitivas [13, 14]. OB3D es una buena alternativa para estudios de reconocimiento de objetos, sin embargo, un objeto consiste en una nube de puntos en el espacio que carece de superficie, color e iluminación, lo que dificulta su aplicación ya que lo que se busca es imitar lo mejor posible las características de un objeto real. Las anteriores son posibilidades tangibles para trabajar el reconocimiento visual y el nombrado de objetos. No obstante, son alternativas que toman bastante tiempo para su adecuación además de requerir de importantes recursos financieros para contar con los equipos necesarios.

El presente proyecto trata de dar una solución a este problema puesto que propone una configuración experimental de bajo costo que utiliza videos de alta definición en lugar de objetos reales. Este sistema permitirá comparar el grado en que la presentación de objetos familiares en video en la pantalla de un computador y la de esos objetos reales, facilitan su reconocimiento y nombrado en contraposición con las representaciones gráficas en dos dimensiones. De esta manera se optimizarían los recursos terapéuticos en la rehabilitación de pacientes con déficit en la percepción visual y en la nominación, reduciendo los problemas derivados de la gestión de objetos reales, y posiblemente ampliando el banco de estímulos visuales terapéuticos. La evidencia disponible no permite conocer la aplicación de la modalidad de presentación de videos como estímulos terapéuticos para la intervención en alteraciones de la percepción visual.

El trabajo interprofesional entre psicólogos, rehabilitadores e ingenieros biomédicos busca aportar alternativas tecnológicas para abordar este problema y facilitar la disponibilidad y accesibilidad de estímulos terapéuticos en la atención de pacientes con alteraciones en la percepción visual.

Este trabajo reporta el diseño, la construcción y el pilotaje de una plataforma de visualización de objetos reales como herramienta para controlar la forma y tiempo de presentación de los estímulos que se utilizarán en una fase experimental subsiguiente con pacientes con agnosia visual. La plataforma permite la visualización de los objetos seleccionados reales en rotación y la generación de los videos de los objetos. Para probar el sistema, se realizó un estudio piloto con personas sanas sin diagnóstico de enfermedad neurológica ni déficit visual, en el que fueron ejecutadas dos tareas: 1) nombrado de objetos presentados en fotografías versus objetos reales y 2) nombrado de objetos reales versus objetos presentados en video.

## MATERIALES Y MÉTODOS

La metodología incluyó cuatro fases. En la Fase 1 se seleccionaron los objetos (estímulos) que serían utilizados en las diferentes modalidades de presentación. La Fase 2 incluyó el diseño y construcción de los dispositivos necesarios para la visualización de estímulos en la modalidad de objetos reales; estos dispositivos consistieron en una plataforma para visualizar los objetos en rotación y en un biombo que los ocultara durante la transición de estímulos en las sesiones experimentales. En la Fase 3 se adquirieron los medios (fotografías y videos) que debían cumplir con ciertos criterios de calidad. Finalmente, en la Fase 4 se realizó una prueba piloto para probar el desempeño del sistema.

## Fase 1 - Selección de estímulos/objetos

Fueron seleccionados 80 objetos del banco de imágenes estandarizado propuesto por Salmon y colaboradores [1], controlando variables como el uso funcional, el agarre, la edad de adquisición, la familiaridad, la complejidad visual, la concordancia en la etiqueta y la frecuencia léxica, obtenidas de las medias y desviaciones estándar reportadas en otros estudios de normatización [2, 15, 16, 17]. Fueron controladas también el uso unimanual o bimanual mediante una tarea de normatización en un grupo de 44 personas típicas, en cuyo caso debían clasificar los 80 objetos en uni o bimanuales, para lo cual se obtuvo del total de estímulos un 58,75% unimanual y 41,25% bimanual, valores que fueron tenidos en cuenta para la distribución homogénea de objetos en las listas de presentación utilizados en los experimentos. Un tratamiento similar se llevó a cabo para la frecuencia de uso de los objetos, en cuyo caso las cinco personas que participarían en la prueba piloto calificaron mediante una escala Likert la periodicidad con la cual hacían uso de los objetos en la lista obteniendo, que de la totalidad de los estímulos, el 42,5% fueron de baja frecuencia y 57,5% de alta. Para el análisis de los resultados de la escala Likert se tuvieron en cuenta los valores de 0 a 10 por cada objeto, clasificando en alta frecuencia a los estímulos con un número mayor o igual a 5 y baja menor de 5.

Los valores de frecuencia léxica de la etiqueta para cada objeto fueron obtenidos del Léxico Informatizado del Español-LEXESP [18] (ver Tabla 1). La frecuencia léxica se corresponde a la aparición de la palabra por un conteo de millón de palabras, de manera que hubo palabras o etiquetas de alta y baja frecuencia. Para los objetos no contenidos en el LEXESP, fueron tomados los valores de Brodeur [2] y Cuetos [15]. Otros objetos (8% del total de estímulos) seleccionados para la prueba piloto no cuentan con valores normalizados.

**Tabla 1.** Listado de los 80 objetos utilizados en el estudio y su valores de frecuencia léxica.

| Objetos            | Frecuencia Léxica | Objetos             | Frecuencia Léxica |
|--------------------|-------------------|---------------------|-------------------|
| alicates           | 23                | armónica            | 2,86              |
| bisturí            | 6,07              | atomizador          | 0,18              |
| bombilla           | 5,36              | batidora            | —                 |
| botella            | 31,61             | bota                | 5,18              |
| bufanda            | 4,11              | cepillo             | 5,18              |
| caja               | 45                | cepillo de dientes  | 39                |
| calculadora        | 2,14              | clip                | 0,54              |
| cámara             | 39,82             | collar              | 6,25              |
| campana            | 11                | compota             | 0,18              |
| cenicero           | 6,43              | control remoto      | 23                |
| cepillo de cabello | —                 | corbata             | 17,14             |
| cigarrillo         | 39,82             | cuchara             | 3,75              |
| cinta métrica      | 0                 | cuchilla de afeitar | 5,58              |
| cinturón           | 14,64             | cuchillo            | 15,36             |
| cortaúñas          | 31                | dados               | —                 |
| depilador          | 30                | descorchador        | —                 |
| destapador         | 24                | destornillador      | 35                |
| encendedor         | 1,61              | escoba              | 2,86              |
| gorra              | 7,68              | esfero              | —                 |
| guante             | 8,39              | gafas               | 33,04             |
| calcetín           | 1,43              | grapadora           | 37                |
| pandereta          | 1,07              | lápiz               | 10                |
| pantalón           | 18,21             | libro               | 194,29            |
| papel higiénico    | 35                | llave               | 22,86             |
| paraguas           | 12,5              | llave alemana       | —                 |
| peine              | 5                 | maracas             | 1,43              |
| pinzas de hielo    | —                 | marcador            | 6,07              |
| pistola            | 26,79             | martillo            | 5,36              |
| pitillo            | 8,93              | pala                | 4,82              |
| plancha            | 5,18              | pelota              | 18,39             |
| rayador            | 15                | perforadora         | 0,54              |
| sacaganchos        | 25                | pincel              | 3,57              |
| salero             | 0,71              | pinza de ropa       | —                 |
| sobre              | 1458,75           | puntilla            | 0,89              |
| taladro            | 1,43              | secador de pelo     | 25                |
| tambor             | 6,79              | sERRUCHO            | 1,07              |
| teléfono           | 80,18             | silbato             | 1,61              |
| tornillo           | 1,43              | sonajero            | 0,71              |
| transportador      | 13                | tajalápiz           | 24                |
| violín             | 5,18              | tijeras             | 4,82              |



**Figura 1.** Algunos de los objetos utilizados.

En aquellos casos en los cuales los valores de las medias de las variables de interés no estaban disponibles en la estandarización en español [15, 16], se tomaron los valores establecidos de la base de datos en inglés [1] para el mismo objeto. Cada sesión experimental quedó constituida por 40 estímulos distribuidos en dos grupos contrabalanceados por las variables de control; de manera que ambos grupos incluyeron objetos de medias equivalentes de uso funcional, agarre, edad de adquisición, familiaridad, complejidad visual, concordancia en la etiqueta y frecuencia léxica. La Figura 1 es una muestra de seis de los 80 estímulos/objetos seleccionados.

## Fase 2 - Diseño y construcción de dispositivos

En esta fase se trabajó en el diseño y construcción de la plataforma de rotación que permitió la toma de videos y fotografías, así como la visualización de objetos en modalidad real. Adicionalmente, se construyó un biombo para ocultar la plataforma y los objetos en modalidad real que fueron presentados a los participantes. Para garantizar el buen funcionamiento de dichos dispositivos y el éxito de las pruebas se establecieron los siguientes requisitos de diseño:

### *Requisitos de diseño*

- Los dispositivos debían ser de bajo costo y portables.
- No debían contener elementos distractores (el participante debe centrar su atención en el objeto).

- La plataforma debía generar el menor ruido posible durante su funcionamiento.
- La velocidad de rotación de la plataforma debía ser ajustable y después permanecer constante.
- El área efectiva de trabajo para colocar los objetos en la plataforma debía ser de  $1200 \text{ cm}^2$  aproximadamente (círculo de  $40 \text{ cm}$  de diámetro).
- La plataforma debía soportar objetos de un peso máximo de 4 kilogramos.
- El biombo debía ser liviano, de fácil manipulación y tener una altura máxima de  $60 \text{ cm}$ .
- El biombo debía tener un área efectiva para ocultar los objetos y la plataforma de  $0,5 \text{ m}^2$ .

Si los dispositivos no cumplían con los requisitos de diseño establecidos estos eran rediseñados o ajustados hasta lograrlo.

Se obtuvo un diseño de la plataforma giratoria como la observada en la Figura 2. Esta requirió una inversión aproximada de 250 dólares americanos. Fue construida en madera aglomerada tipo MDF (Medium Density Fibreboard) y acrílico blanco de 6 milímetros de grosor. El sistema de rotación de la base (círculo blanco) es por contacto directo con un motor con motoreductor de 12 voltios. El diagrama de bloques de la plataforma puede apreciarse en la Figura 3. Este consta de un microcontrolador (PIC12C671) en donde una entrada analógica modifica el ancho de pulso (PWM por sus siglas en inglés) que afectará la velocidad del motor. El motor posee un motoreductor que transmite su movimiento directamente a la base de rotación. Por último, un regulador de corriente alimenta una cinta de LEDs que ilumina la base de la plataforma para resaltar los objetos allí colocados.





Figura 2. Plataforma desarrollada.

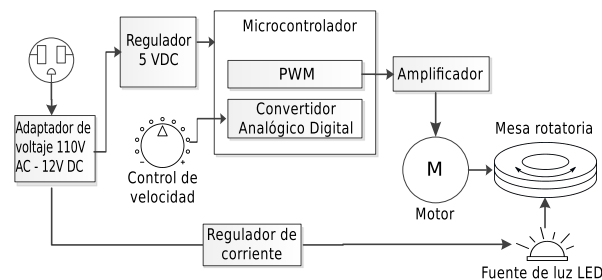


Figura 3. Diagrama de bloques de la plataforma.

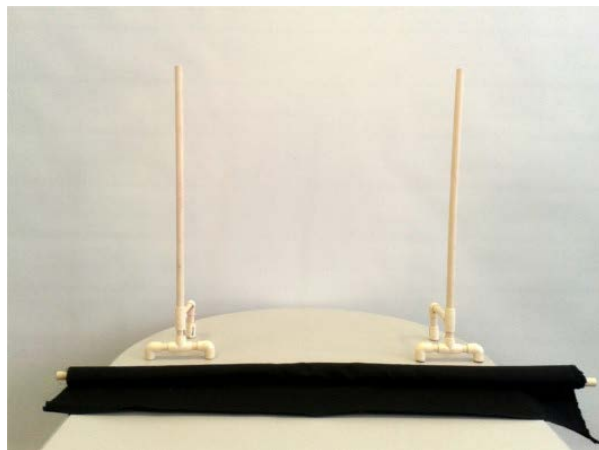


Figura 4. Biombo construido.

La parte de madera de la plataforma se pintó de color negro mate para evitar reflejos y para resaltar los objetos ubicados en la base blanca.

El biombo consiste en dos estructuras verticales simples construidas en tubo PVC (Policloruro de vinilo) de media pulgada (Figura 4). A estas dos estructuras verticales se acopla un telón negro mediante un tubo horizontal hecho también de PVC (no acoplado en la Figura 4). El biombo se

arma y se desarma rápidamente ocupando un mínimo de espacio. Requirió una inversión de aproximada de 30 dólares americanos.

### Fase 3 - Adquisición de medios

En esta fase fue necesario adquirir los medios (fotografías y videos) con la mejor calidad posible, es decir, que tanto las fotografías (imágenes) como los videos debían ofrecer la máxima información posible al observador acerca del objeto, tal y como si se tratara de la observación del objeto real.

En un inicio, las fotografías se tomaron sin contar con el factor facilitador de la manipulación, es decir sin la condición de ubicar los objetos en una posición que estuvieran “listos para usarse”; también se tomaron las fotografías con una resolución de 1334 X 750 píxeles (media resolución) para evitar el elevado peso en MB (megabytes) de las fotografías finales, sin embargo, objetos pequeños como un depilador presentaron pixelación. Por otro lado, los videos presentaron una irregularidad porque la rotación de la plataforma no había sido controlada, lo que produjo una inestabilidad visual en la imagen del video.

Con estas anomalías, se reinició el proceso de toma de fotos y videos pero ahora en alta resolución y en alta definición respectivamente para lo cual se debía cumplir con los siguientes criterios de la calidad:

#### *Criterios de calidad de los medios*

- Las fotografías debían ser tomadas en alta resolución (2592 X 1728 píxeles) a 14 bits en grano fino.
- Los videos debían ser tomados en Full HD (1920 X 1080 píxeles).
- Los medios debían tener las mismas dimensiones entre sí, es decir, el mismo tamaño todas las fotografías y el mismo tamaño todos los videos.
- La distancia entre la cámara y los objetos a fotografiar/filmar debía ser siempre constante, 80 *cm*.



**Figura 5.** Configuración para la toma de videos y fotografías.

- El ángulo formado entre el plano horizontal de la plataforma y la dirección de enfoque de la cámara debía ser el mismo en la toma de fotos y de videos,  $30^\circ$ . Esto corresponde aproximadamente al ángulo de visión de una persona sentada que mira hacia la plataforma.
- Las mismas condiciones de iluminación para la toma de fotografías y de videos.
- Posición adecuada para manipulación del objeto sobre la plataforma en la fotografía y en el primer fotograma del video (véase la posición de los objetos en la Figura 1).
- La grabación de los videos debe tener estabilización de imagen y sin pista de audio.

Los medios fueron auditados por dos investigadores en forma independiente para conceptualizar la calidad de los mismos, de no coincidir con la evaluación, estos debían ser repetidos hasta lograrlo. La configuración final utilizada para la toma de fotografías y videos se observa en la Figura 5 (el objeto en este caso es un guante). Se utilizó una cámara profesional Canon EOS Rebel T2i de 18 mega píxeles con un flash inalámbrico marca Vivitar, trípodes y accesorios de mesa. El monitor utilizado para la visualización de los medios fue de 20 pulgadas referencia HP LV2011, con retroiluminación LED y resolución 1920X1080 a 60Hz.

Al final de esta etapa y después de varias repeticiones, los medios incluían 80 fotografías de alta resolución y 80 videos en alta definición.

#### Fase 4 - Prueba piloto

##### *Participantes*

Se escogieron por conveniencia cinco participantes adultos típicos, (4 mujeres y un hombre) sin compromisos cognitivos ni viso-perceptuales, con una media de edad de 54 años y con un nivel de escolaridad que incluía secundaria, técnico y profesional. Una vez leído y firmado el consentimiento informado por parte de los participantes, se aplicaron como pruebas de tamizaje el test del Estado Mínimo Mental [19] (funciones mentales superiores), el test de Boston (dominio del lenguaje) [20] y el Visual Motor Integration (subtest de percepción visual) [21]. Los resultados de la aplicación de estas pruebas determinaron la continuidad de los participantes, por cuanto descartaban la existencia de alteraciones cognitivas, visuales y del lenguaje.

La Figura 6 esquematiza la prueba piloto realizada, la cual tuvo las siguientes características:



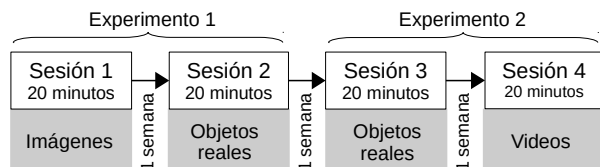


Figura 6. Metodología de la prueba piloto.

- Se diseñaron dos experimentos.
- En cada experimento, cada participante asistió a dos sesiones de 20 minutos cada una, en dos días distintos con un intervalo de una semana entre una sesión y otra. El ambiente donde se realizó la tarea fue controlado, garantizando luminosidad, comodidad y ausencia de interferencia para la tarea.
- El Experimento 1 condujo dos sesiones: una en donde fueron presentados 40 estímulos en fotografía, y una semana después una sesión en la cual se le presentaron los mismos 40 objetos pero en versión real, rotando en la plataforma.
- En el Experimento 2 se repitió la misma metodología del Experimento 1 pero, en este caso, la presentación de objetos fue en versión real y en video; además, se utilizaron 40 objetos diferentes. En total, la prueba piloto usó 80 estímulos.
- La presentación de los estímulos en fotografías y videos se hizo por medio del programa informático SuperLab<sup>1</sup>, del cual el grupo de investigación posee una licencia para su uso; sin embargo, como la filosofía del grupo es fomentar también el uso de tecnologías de bajo costo, en una etapa posterior del proyecto se utilizarán herramientas gratuitas como OpenSesame<sup>2</sup> o como PsychoPy<sup>3</sup>. Estas son herramientas de código abierto y permiten la presentación de estímulos y la recogida de datos para una amplia gama de experimentos en neurociencia y psicología [22, 23]. En el caso de SuperLab, es una herramienta que facilita la presentación de estímulos para la investigación comportamental en neurociencias cognitivas. Con este programa es posible temporizar y controlar el orden de presentación de los estímulos. Al momento de presentar un estímulo en fotografía o en video, el sistema generaba un discreto sonido (tipo *bip*) y se visualizaba dicho estímulo por un tiempo máximo de ocho segundos para luego pasar al siguiente estímulo; pero si el sujeto respondía antes de este intervalo de tiempo, el sistema permitía avanzar al siguiente estímulo oprimiendo una tecla. El tiempo de visualización de cada objeto fue obtenido de la descripción de procedimientos similares en estudios previos de reconocimiento de objetos usando este tipo de herramientas [24, 25].
- La totalidad de las sesiones fueron registradas mediante grabaciones de audio y analizadas con el programa informático Praat<sup>4</sup>, que es una herramienta para estudios científicos de lingüística, que se encarga de analizar el espectrograma de sonidos grabados, lee sonidos registrados directamente con el programa y los edita para comprobar la entonación, intensidad, volumen de voz y otros aspectos más complejos del habla. En este caso se usó para controlar el tiempo de inicio y finalización de nombrado de cada objeto.
- Para un análisis más exhaustivo, Praat es capaz de analizar fragmentos

<sup>1</sup>SuperLab, Cedrus Corporation, <http://www.superlab.com>.

<sup>2</sup>OpenSesame, Herramienta para psicología experimental y neurociencia, <http://osdoc.cogsci.nl/>

<sup>3</sup>PsychoPy, The University of Nottingham, <http://www.psychopy.org/>

<sup>4</sup>Praat, Boersma Paul y Weenink David, <http://www.fon.hum.uva.nl/praat/>.

de sonido, visualizar las formas de onda y sus componentes frecuenciales. En dicho análisis, se obtuvieron las latencias en milisegundos entre el momento de presentación de los estímulos (ubicando el estímulo sonoro *bip*) y la producción sonora del sujeto al nombrar dicho estímulo. Cuando la presentación de los estímulos era en versión real rotando en la plataforma, las latencias también se obtuvieron de las grabaciones de audio pero en este caso la señal sonora tipo *bip*, fue reemplazada por la emisión oral /¡ya!/ generada por el terapeuta al momento de soltar el manto negro del biombo. A partir de este momento se despreciaban los siguientes 380 milisegundos que era el tiempo que le tomaba al manto negro caer del biombo. De esta manera se garantizaba que el tiempo cero de latencia comenzaba cuando el objeto era totalmente descubierto por el manto negro del biombo.

La Figura 7 muestra en su parte izquierda la situación inicial vista por un sujeto en una sesión de estímulos reales, y en la parte derecha, la situación vista por el mismo sujeto 380 milisegundos después de haber soltado el manto negro del biombo.

Fueron obtenidos los tiempos de reacción o latencias de nombrado de los objetos, representados en el tiempo en milisegundos que tardaba el participante en pronunciar el nombre de un objeto después de que este ha sido mostrado. Luego se analizaron los datos por medio del programa estadístico R<sup>5</sup>. Ya que la distribución de los datos (latencias) fue no normal, se aplicó la prueba estadística no paramétrica de Wilcoxon para medir las diferencias entre las latencias de nombrado para las condiciones de presentación de las sesiones del Experimento 1 (imágenes versus objetos reales) y del Experimento 2 (videos versus objetos reales) [26].



**Figura 7.** Montaje ejemplo para sesiones con estímulos reales.

## RESULTADOS

Los resultados relacionados con la plataforma y el biombo como prototipos finales pueden apreciarse en las Figuras 2 y 4 respectivamente.

Considerando que la distribución de los datos no fue normal para ninguna de las sesiones del estudio (empleando el test de normalidad de Shapiro-Wilk), se obtuvieron las medianas y los rangos intercuartílicos y se utilizó la prueba de Wilcoxon (Alpha=0,05; una cola) para establecer si existían diferencias estadísticamente significativas entre las medianas de las latencias [26].

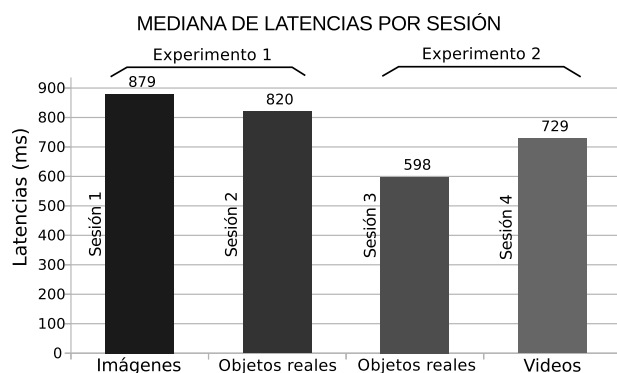
Las medianas y los rangos intercuartílicos (RI) de los tiempos de reacción (en milisegundos) para el conjunto de las cuatro sesiones fueron: sesión 1-imágenes (mediana=879 - RI=445), sesión 2-Objetos reales (mediana=820 - RI=769), sesión 3-objetos reales (mediana=598 - RI=564) y sesión 4-videos (mediana=729 - RI=764), estos datos se observan en la Figura 8. Por su parte, la Figura 9 muestra las medianas de las latencias obtenidas por modalidad de presentación, es decir: imágenes (sesión 1), objetos reales (sesiones 2 y 3) y videos (sesión 4).

La comparación se hizo entre las latencias registradas en las tres modalidades de presentación: imágenes versus objetos reales unidos (sesiones 2 y 3), videos versus objetos reales unidos (sesiones 2 y 3) e imágenes versus videos. Los resultados se muestran en la Tabla 2 y los valores con asterisco (\*)

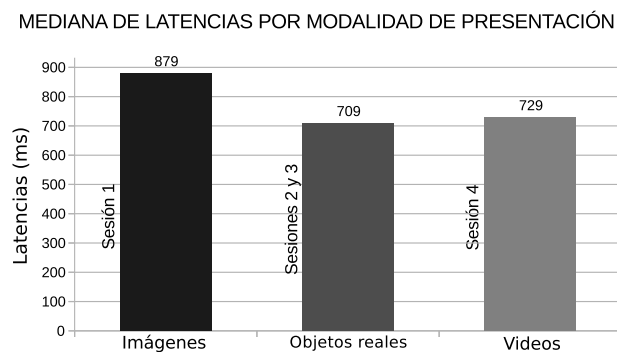
<sup>5</sup>The R Project for Statistical Computing, <http://www.r-project.org/>

representan significancia estadística.

Analizando las modalidades de presentación (Figura 9), se confirma la hipótesis establecida: los tiempos de reacción para la condición de objetos en imágenes fueron mayores que los registrados para el reconocimiento y nombrado de los objetos reales y los videos. No hubo mayores diferencias entre las modalidades de presentación en videos y objeto real. Sin embargo, aunque los tiempos de reacción fueron mayores para la condición de imágenes en comparación con las otras dos modalidades de presentación visual, contrario a lo esperado, no se encontraron diferencias estadísticamente significativas entre los tiempos de reacción en las latencias de las modalidades de imagen y video.



**Figura 8.** Resultados de la prueba piloto por sesiones.



**Figura 9.** Resultados de la prueba piloto por modalidad de presentación.

**Tabla 2.** Resultados estadísticos.

| Test de Wilcoxon (pareado, una cola, alpha 0.05) | p-valor   |
|--|-----------|
| Imágenes vs Objetos reales unidos                | 0.005832* |
| Videos vs Objetos reales unidos                  | 0.4598    |
| Imágenes vs Videos                               | 0.4125    |

## DISCUSIÓN

El estudio del reconocimiento de objetos es central en la investigación básica y clínica de las funciones cognitivas. Sin embargo, esta área de la investigación enfrenta dificultades a la hora de encontrar recursos tecnológicos asequibles que permitan la visualización y el reconocimiento de objetos estandarizados o de selección arbitraria por parte de terapeutas e investigadores. En este sentido, son valiosos los aportes que buscan ofrecer una alternativa de bajo costo para la visualización y el reconocimiento de objetos que pueda ser utilizada en la práctica clínica.

Los resultados del presente trabajo apoyan la dependencia en la modalidad de presentación en la percepción y nombrado de objetos, dado que los tiempos de reacción para la condición de objetos reales fueron menores que para las otras dos modalidades de presentación. En este sentido, al parecer las propiedades de volumen, permiten la exploración de características de los objetos que las fotografías no tendrían; razón por la cual, los tiempos de reacción para los objetos reales y videos fueron menores que la condición en imágenes, produciéndose un efecto de facilitación para su identificación.

La no significancia estadística de las diferencias mostradas entre las condiciones de imagen y video quizás pueda ser explicada porque los datos fueron obtenidos de personas que no presentan ningún tipo de fallo en la percepción visual y, por tanto, dichas diferencias en la modalidad de presentación no resulten tan evidentes en casos típicos. Será necesario comparar estos resultados con el desempeño de un grupo atípico como en el caso de pacientes con agnosia visual.

Las personas reconocemos una amplia gama de objetos con poco esfuerzo, a pesar del hecho que el objeto puede variar en diferentes puntos de vista, tamaños o escala e incluso cuando están rotados. Mucho se ha discutido sobre si el reconocimiento de objetos es un asunto de la percepción de la forma en comparación con la congruencia de la apariencia [6]. La apariencia de un objeto en una imagen depende de condiciones relacionadas con el registro fotográfico como la forma y tamaño del objeto, de las propiedades refractarias de la luz, del lugar que ocupan o del contexto físico en el cual se presente y hasta de las condiciones de iluminación [14]. Unas condiciones intrínsecas e invariables de los objetos como su forma, y otras, extrínsecas y variables como la posición. Ello hace que para que el sistema visual sea hábil para reconocer objetos estos deban tener un modelo representado en la memoria. Es claro que la apariencia de los objetos juega un papel primordial en el aprendizaje, la evocación y el reconocimiento visuales. Poco sabemos sobre cómo nuestro cerebro construye estas representaciones exactas de los objetos. Lo que a hoy quizás parece ser aceptado es que el cerebro almacena propiedades constantes de los objetos y ello facilita su reconocimiento. Para el caso de la prueba piloto realizada en el presente estudio se usaron objetos cuya apariencia estaba ligada a una misma categoría, herramientas, que implican una relación con su uso funcional vinculado con un programa motor de activación cerebral. Ello podría explicar la no diferencia en tiempo de reacción cuando se analizan los objetos en función de su uso, es decir, que al tratarse de herramientas hay una praxia asociada que facilita su evocación, por la activación de la representación mental del patrón motor. Habría que explorarse otro tipo de objetos no funcionales para estimar el valor predictivo de esta variable relacionada con la manipulabilidad. Una limitación en esta prueba piloto fue el 8% de objetos

seleccionados que no contaron con valores normalizados.

## CONCLUSIONES

Todos los campos científicos aplicados, como la rehabilitación, dependen de la investigación para avanzar en la provisión e innovación de tratamientos, servicios y productos tecnológicos. Así mismo, el trabajo de equipos interprofesionales (ingenieros, psicólogos y terapeutas) en la investigación en rehabilitación garantiza que la tecnología propuesta se aproxime más a las necesidades y realidades de la población objetivo. Por ello, es necesario avanzar en la investigación inter y multidisciplinar que garantice soluciones integrales de bajo costo pero de alto impacto social.

Tal como la investigación en percepción ha mostrado, la modalidad de presentación de los estímulos visuales determina su probabilidad de reconocimiento y nombrado. Los resultados de la prueba piloto del presente estudio van en la misma dirección. Una segunda fase de este estudio deberá responder a qué tipo de características de los objetos resulta clave para su identificación y qué correlación existe entre la modalidad de presentación y las variables intrínsecas de los objetos como tamaño, forma, complejidad visual, manipulabilidad, frecuencia de uso y frecuencia léxica, edad de adquisición, entre otras. De igual forma, deberá avanzarse en cómo estas variables de los estímulos visuales afectan distintamente a quienes tienen alteraciones en la percepción visual, como en el caso de la agnosia asociativa. Debe reconocerse, como limitaciones de este tipo de estudios, la dificultad para controlar las características propias de los sujetos que implica diversidad en las estrategias de resolución de tareas perceptivas-cognitivas, los factores intrínsecos relacionados con la experiencia previa en el uso de los objetos, el nivel socioeducativo, la edad y el género. Los desarrollos, la disponibilidad y aplicabilidad de dispositivos tecnológicos que mediaron

el diseño y conducción del estudio en los aspectos técnicos tuvieron una implicación en el alcance de la investigación.

### Agradecimientos

Los autores de este trabajo quieren agradecer a los estudiantes: Álvaro José Bocanegra Pérez, Sergio David Sierra Marín, Juan Sebastián Lara Ramirez y Paola Andrea Cruz Vargas, por su apoyo y colaboración en la realización de esta investigación; a Hernán Alfredo Bernal Muñoz por su apoyo técnico en fotografía y diseño y a Adriana María Ríos Rincón por su apoyo en el análisis estadístico.

### REFERENCIAS

- [1] J. P. Salmon, P. a. McMullen, and J. H. Filliter, “Norms for two types of manipulability (graspability and functional usage), familiarity, and age of acquisition for 320 photographs of objects.,” *Behavior research methods*, vol. 42, pp. 82–95, feb 2010.
- [2] M. B. Brodeur, K. Guérard, and M. Bouras, “Bank of Standardized Stimuli (BOSS) Phase II: 930 New Normative Photos,” *PLoS ONE*, vol. 9, p. e106953, jan 2014.
- [3] C. A. Ankerstein, R. A. Varley, and P. E. Cowell, “A new set of 84 colored nonobjects.,” *Behavior research methods*, vol. 41, pp. 172–6, feb 2009.
- [4] M. J. Farah, *Visual Agnosia: Disorders of Object Recognition and What They Tell Us About Normal Vision*. Cambridge, MA.: MIT Press, 1990.
- [5] O. H. Turnbull, J. Driver, and R. A. McCarthy, “2D but not 3D: pictorial-depth deficits in a case of visual agnosia.,” *Cortex; a journal devoted to the study of the nervous system and behavior*, vol. 40, pp. 723–38, jan 2004.
- [6] H. Chainay and G. W. Humphreys, “The real-object advantage in agnosia: Evidence for a role of surface and depth information in object recognition.,” *Cognitive neuropsychology*, vol. 18, pp. 175–91, mar 2001.
- [7] E. De Renzi, “Disorders of visual recognition.,” *Seminars in neurology*, vol. 20, pp. 479–85, jan 2000.
- [8] D. A. Wolk, H. B. Coslett, and G. Glosser, “The role of sensory-motor information in object recognition: evidence from category-specific visual agnosia.,” *Brain and language*, vol. 94, pp. 131–46, aug 2005.
- [9] M. N. Magnié, M. Besson, M. Poncet, and C. Dolisi, “The Snodgrass and Vanderwart set revisited: norms for object manipulability and for pictorial ambiguity of objects, chimeric objects, and nonobjects.,” *Journal of clinical and experimental neuropsychology*, vol. 25, pp. 521–60, jun 2003.
- [10] S. Zhang, “Recent progresses on real-time 3D shape measurement using digital fringe projection techniques,” *Optics and Lasers in Engineering*, vol. 48, pp. 149–158, feb 2010.
- [11] S. Khalfaoui, R. Seulin, Y. Fougerolle, and D. Fofi, “An efficient method for fully automatic 3D digitization of unknown objects,” *Computers in Industry*, vol. 64, pp. 1152–1160, dec 2013.
- [12] Z. Liu, H. Qin, S. Bu, M. Yan, J. Huang, X. Tang, and J. Han, “3D real human reconstruction via multiple low-cost depth cameras,” *Signal Processing*, vol. 112, pp. 162–179, jul 2015.
- [13] “[OB3D] A data base of 3 dimensional Objects.”
- [14] S. Buffat, V. Chastres, A. Bichot, D. Rider, F. Benmussa, and



- J. Lorenceau, "OB3D, a new set of 3D objects available for research: a web-based study.," *Frontiers in psychology*, vol. 5, p. 1062, jan 2014.
- [15] F. Cuetos, A. W. Ellis, and B. Alvarez, "Naming times for the Snodgrass and Vanderwart pictures in Spanish," *Behavior Research Methods, Instruments, & Computers*, vol. 31, pp. 650–658, dec 1999.
- [16] M. C. Sanfeliu and A. Fernandez, "A set of 254 Snodgrass-Vanderwart pictures standardized for Spanish: Norms for name agreement, image agreement, familiarity, and visual complexity," *Behavior Research Methods, Instruments, & Computers*, vol. 28, pp. 537–555, dec 1996.
- [17] M. B. Brodeur, E. Dionne-Dostie, T. Montreuil, and M. Lepage, "The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research.," *PloS one*, vol. 5, p. e10773, jan 2010.
- [18] N. Sebastián-Gallés, *LEXESP: léxico informatizado del español*. Edicions de la Universitat de Barcelona, 2000.
- [19] M. F. Folstein, S. E. Folstein, and P. R. McHugh, "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician.," *Journal of psychiatric research*, vol. 12, pp. 189–98, nov 1975.
- [20] H. Goodglass, E. Kaplan, and B. Barresi, *Test de Boston para el diagnóstico de la afasia: adaptación española*. Bogotá: Editorial Medica Panamericana, 3 ed. ed., 2005.
- [21] K. Beery and N. Buktenica, *VMI, the Developmental Test of Visual-motor Integration: Specimen Set*. Modern Curriculum Press, 1982.
- [22] S. Mathôt, D. Schreij, and J. Theeuwes, "OpenSesame: an open-source, graphical experiment builder for the social sciences.," *Behavior research methods*, vol. 44, pp. 314–24, jun 2012.
- [23] J. W. Peirce, "PsychoPy—Psychophysics software in Python.," *Journal of neuroscience methods*, vol. 162, pp. 8–13, may 2007.
- [24] F. J. Moreno-Martínez and P. R. Montoro, "An ecological alternative to Snodgrass & Vanderwart: 360 high quality colour images with norms for seven psycholinguistic variables.," *PloS one*, vol. 7, p. e37527, jan 2012.
- [25] D. D. Leeds, J. A. Pyles, and M. J. Tarr, "Exploration of complex visual feature spaces for object perception.," *Frontiers in computational neuroscience*, vol. 8, p. 106, jan 2014.
- [26] L. G. Portney and M. P. Watkins, *Foundations of Clinical Research: Applications to Practice. Chapter 22*. Pearson/Prentice Hall, 2009.

