

## Análisis ROC y z-ROC en memoria de reconocimiento: modelos continuos, de umbral y asimétricos

Manuel Pelegrina del Río<sup>1</sup>, Enrique Emberley Moreno<sup>2</sup>, Agustín Wallace Ruiz<sup>1</sup> y Rafael Marín Lamela<sup>1</sup>

<sup>1</sup> Universidad de Málaga y <sup>2</sup> Centro de Profesores Campo de Gibraltar

En esta investigación aplicamos la característica operativa del receptor (ROC) para evaluar la memoria de reconocimiento de ítems mediante palabras de clase abierta (CA) y clase cerrada (CC). El objetivo consistió en evaluar procesos y modelos de la memoria a partir del tipo de estímulos utilizados: palabras en lengua española, clasificadas según características gramaticales. En el registro de datos se aplicó una escala de estimación (EE) con el objeto de obtener los niveles apropiados de respuesta. Se obtuvieron diferencias estadísticamente significativas entre el procesamiento de ambos tipos de palabras. Obtuvimos modelos de umbral, continuos y de varianza desigual. Con ello pudimos contribuir a la resolución de la controversia actual sobre los modelos más adecuados para evaluar la memoria de reconocimiento de ítems. Finalmente, se argumenta sobre la diferenciación e integración de los modelos desde un punto de vista empírico.

*ROC and z-ROC analysis in recognition memory: Continuous, threshold, and asymmetric models.* In this research, we applied the operating characteristic of the receiver (ROC) to evaluate item-recognition memory by means of open class words (CA) and closed class words (CC). The purpose was to evaluate memory processes and models, according to the type of stimuli used: words in Spanish, classified according to grammar characteristics. To register the data, a rating scale was applied to obtain the appropriate response levels. Statistically significant differences between the processing of both types of words were obtained. We evaluated models of threshold, continuous, and unequal variance. As a result, we could contribute to the resolution of the present controversy about the models adapted to evaluate the memory of item recognition. Lastly, the differentiation and integration of the models from an empirical point of view is discussed.

Una cita histórica obligada nos indica que Egan (1958) fue el primer investigador que, mediante un trabajo muy citado, incorpora la palabra como señal en el ámbito de la teoría de la detección de señales (TDS). Los métodos de investigación en memoria de reconocimiento en el ámbito ROC fueron ampliamente desarrollados por Egan (1975).

En este sentido hay líneas diferenciadas de investigación abiertas. Una de ellas, de naturaleza integradora, se ha centrado en si es más adecuado un modelo de umbral en el proceso de *recollection* (recuerdo) juntamente con un modelo de reconocimiento por familiaridad (Mickes, Johnson y Wixted, 2010). La segunda línea de investigación, también integradora, parte de un modelo basado en la TDS: un modelo de desigual varianza que asume un único proceso continuo de la fuerza de los ítems en memoria, en el que los ítems «viejos» (ítems señal frente a ítems ruido) se reconocen según distribuciones con diferentes varianzas (Jang, Wixted y Huber, 2009; Pratte, Jeffrey y Morey, 2010).

Nuestro objetivo general consistió en hacer una réplica con ítems en lengua castellana utilizando dos categorías de palabras, CA y

CC. La razón de porqué utilizamos palabras clasificadas según su clase se debe a que éstas maximizan las predicciones en cuanto a la variabilidad de las curvas ROC obtenidas, frente a la clasificación basada solo en la frecuencia. Ello se debe a resultados empíricos previos (Pelegrina y Tejeiro, 2006), y a que la CA y CC son palabras con contenido semántico versus conectores (sin significado propio), respectivamente. Son categorías cualitativamente distintas (y extremas en su significado). Supuestamente en las palabras de CA intervienen más procesos cognitivos y en las de CC intervienen más procesos asociativos. Ello además nos permite aplicar el análisis ROC a las estructuras gramaticales del lenguaje.

Las palabras de CA: nombres y adjetivos tienen mayor riqueza semántica que las palabras de CC: preposiciones, conjunciones y algunos adverbios. La utilización de este tipo de palabras encaja en nuestros objetivos, ya que ambos tipos de palabras predicen curvas ROC diferenciadas para los modelos que proponemos en esta investigación. En efecto, este enfoque nos permitirá evaluar las posibles asimetrías según la forma de respuesta de los participantes con relación al tipo de ítems utilizados y en la línea de Pratte, Ruoder y Morey (2010) ya iniciada anteriormente por Glanzer, Kim y Hilford (1999).

La investigación del procesamiento de palabras según su clase se inicia a partir de la década de los setenta. Los trabajos iniciales asumen que hay diferentes conjuntos de palabras en el léxico mental y que su procesamiento es computacionalmente distinto

---

Fecha recepción: 21-12-10 • Fecha aceptación: 19-4-11

Correspondencia: Manuel Pelegrina del Río

Facultad de Psicología

Universidad de Málaga

27071 Málaga (Spain)

e-mail: pelegrina@uma.es

(Biassou, Nespoulous, Dordain y Harris, 1977). A este tipo de procesamiento se la ha denominado «hipótesis de la doble ruta». Los trabajos iniciales se centraron en la afasia (Bradley, 1978; Gordon y Caramazza, 1983). Pero en las personas sin afasia se ha detectado un funcionamiento diferencial de las palabras de clase cerrada (Friederici, 1985). Además, en personas agramáticas se ha detectado que es más probable que omitan palabras de CC que de CA (Biassou et al., 1997).

Por otra parte, en relación a los parámetros ROC y z-ROC se han obtenido diferencias importantes entre palabras de CA y CC (Pelegrina y Tejeiro, 2006). En un segundo experimento de este trabajo se incrementó el tiempo de exposición de la palabra (de 250 ms de exposición se pasó a 450 ms). En él destacamos la reducción de las falsas alarmas casi a la mitad en palabras de CA. Sin embargo, las palabras de CC destacan por el aumento considerable en aciertos.

Lo anterior puede explicarse porque las palabras de CA, al aumentar en exposición, aumentan el umbral de aprendizaje significativo, decreciendo los errores; mientras que en palabras de CC al aumentar el tiempo de exposición aumenta la posibilidad de memorización por asociación. Pero si analizamos los resultados en afásicos de Broca observamos que los errores se producen en elementos fonológicos de palabras de CC (Biassou, Obler, Nespoulous, Dordain y Harris, 1997) y no de CA. Lo que nos permite concluir que hay al menos dos formas diferentes de memorizar cada clase de palabras, tanto en participantes normales como en agramáticos y afásicos.

Además, hemos de añadir que las palabras clasificadas según su clase incorporan una variable de confundido, nos referimos a la frecuencia de uso, que resulta ser diferente en palabras de CA y CC. Así, por ejemplo, en lengua alemana las palabras de CA son muy poco frecuentes comparadas con las palabras de CC. Eso indica que el procesamiento de las palabras según su clase está determinado en parte por la variable frecuencia de uso (véase Münte et al., 2001), véanse también otros ejemplos similares en Bröder y Schüz (2009) y Gorman (1961). En lengua castellana ocurre algo parecido, como mostraremos en el apartado de «material». Por ello, para neutralizar el efecto de frecuencia en las listas de ítems-palabras de CA y CC, organizamos un diseño en el que se balanceó la frecuencia de uso dentro de cada clase.

Considerando todo lo anterior, y mediante palabras de CA y CC, nos planteamos si son más adecuados los modelos basados en la teoría de la detección de señales (TDS) o, por el contrario, los modelos de umbral. Esta polémica, a pesar de su actualidad, ya fue tratada por Coombs, Dawes y Twersky (1970), pero aún no está resuelta. Para ello, diferentes investigadores optan por modelos que van desde la TDS (Glanzer, Kim y Hilford, 1999) hasta autores defensores del umbral, como Rouder y Morey (2009), o bien aquellos que optan por dos modelos separados (Yonelinas y Park, 2007) pasando por los modelos mixtos de carácter integrador, como Wixted (2007) y Jang Wixted y Huber (2009).

Así, pretendemos alcanzar los objetivos concretos siguientes: 1. Comprobar si los participantes evalúan de forma diferencial los distintos niveles de la escala de estimación aplicada, evaluable mediante la tendencia de las curvas ROC. 2. Utilizar ítems-palabras, aplicando criterios de clasificación estadísticos y gramaticales junto con la clasificación basada en frecuencia de uso. 3. Registrar los datos mediante TDS y representarlos mediante curvas ROC y z-ROC con objeto de observar la continuidad de las respuestas según la TDS o la aparición de umbrales de respuesta de acuerdo con las teorías del umbral. 4. Registrar la posible asimetría de los resultados respecto a la diagonal negativa de las curvas ROC.

Además, pretendemos comprobar si hay un solo nivel de respuestas en el ámbito ROC. En este caso, los datos se explicarían mediante la TDS clásica; por el contrario, si hubiera dos o más niveles de respuesta los datos se explicarían o bien mediante modelos de umbral, o bien mediante un modelo de detección de señales asimétrico o de diferente varianza (Jang, Wixted y Huber, 2009) que denominamos «modelo integrador de la TDS». Inicialmente los modelos de umbral fueron propuestos por Blackwell (1963) y los modelos de la TDS por Green y Swets (1966); véase además MacMillan y Creelman (1990) (figura 1).

Figura 1A

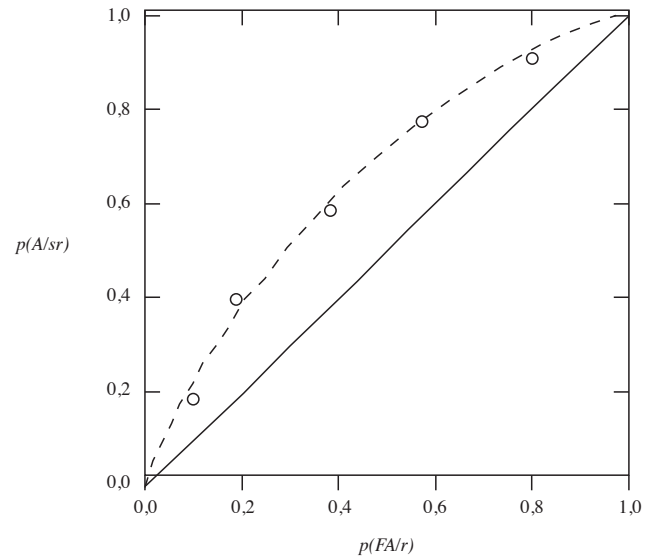
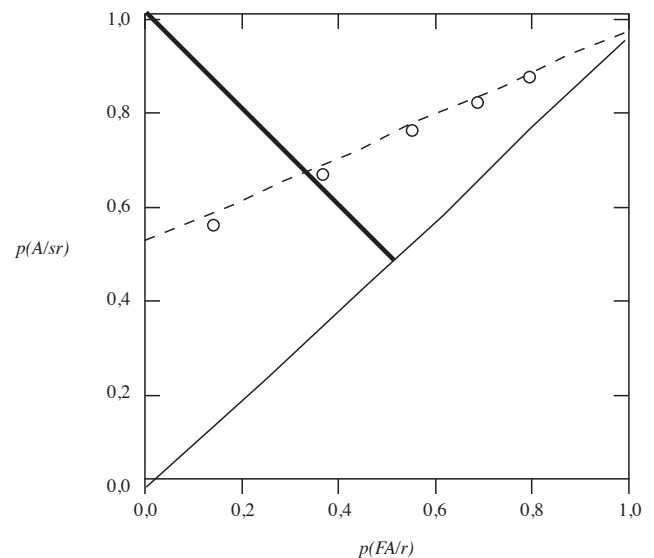


Figura 1B



**Figura 1.** A y B. 1A: curva ROC teórica propia de la TDS. B: curva ROC teórica propia de los modelos de umbral alto. La línea discontinua representa la curva ROC en la figura 1A y el umbral en la 1B. La línea continua de color negro resaltado representa la diagonal negativa sobre la que se mide la asimetría o desigualdad de varianza

Suponemos que hay un continuo en la obtención del umbral en memoria de reconocimiento de ítems si utilizamos palabras de CC, que producirán curvas ROC continuas, de acuerdo con la TDS. Mientras que en palabras de CA, mediante un tiempo de procesamiento mayor de 300 msec, se produciría umbral; o bien diferentes varianzas (o asimetrías) si apelamos a la TDS.

Para conseguir estos objetivos específicos se incluyen palabras clasificadas según su clase (variable cualitativa) y el control de su frecuencia de uso (variable cuantitativa). La clase de palabra es dicotómica (clase abierta y cerrada). Lo anterior representa un diseño con dos niveles en la clasificación de las palabras: CA y CC. Se obtuvieron dos formas de respuesta correspondientes a los aciertos y falsas alarmas, y cinco niveles correspondientes a la formas de respuesta mediante la EE.

#### Método

##### Participantes

Participaron 52 estudiantes (19 hombres y 33 mujeres) entre 19 y 27 años ( $M = 23.3$  y  $DT = 1.258$ ). Todos ellos sin patología manifiesta y con una visión normal o corregida.

##### Instrumentos

En primer lugar establecimos una base de ítems compuesto por una muestra de 80 palabras basada en la frecuencia de uso (FA y FB) y la clase (clase abierta y cerrada). La frecuencia de las palabras de CA (nombres y adjetivos) se situó entre 44 y 696 en frecuencia de uso, véase Juiland y Chang-Rodríguez (1964) y entre 53 y 1.798 en Alameda y Cuetos (1995). La media de la CA se situó entre 48,5 y 1.247 de frecuencia ( $\bar{X} = 647,75$ ). La frecuencia de uso en palabras de CC se situó entre 21 y 2.265 en Juiland y Chang-Rodríguez (1964) y entre 35 y 1.271 en el diccionario de Alameda y Cuetos (1995). La media en frecuencia de palabras de CC se estimó entre 28 y 1.765 ( $\bar{X} = 911,5$ ). Como dijimos anteriormente, observamos un desequilibrio en relación a la frecuencia (véase Münte et al., 2001). Este desequilibrio favorece el parecido entre palabras de clase cerrada y frecuencia alta, actuando la frecuencia como posible variable de confundido. Esto ha quedado solucionado parcialmente al confirmarse que el efecto de clase es una variable que maximiza la varianza sistemática primaria frente a la variable frecuencia de uso (Pelegrina y Tejeiro, 2006).

Se eligieron las 80 palabras que mejor cumplían los requisitos para el balanceo según la frecuencia (tabla 1). Constaban de 1 o 2 sílabas y entre 3 y 5 letras. Éstas se organizaron en dos grupos, uno constaba de 40 palabras de CA, de las cuales 20 eran de FA y 20 de FB. El otro grupo estaba formado por otras 40 palabras de CC de las cuales 20 eran de FA y 20 de FB, tal y como hemos detallado en el párrafo anterior. Las dos listas de palabras fueron presentadas mediante el programa E-prime (Schneider, Eschman y Zuccolotto, 2002). Hemos de aclarar que desde la década de 1960 se utilizan listas de palabras y/o sílabas en la TDS (véanse trabajos de estas fechas citados en el presente artículo). Así, por ejemplo, Baddeley (1983, p. 173) propone como ejemplo típico una lista de 15 ítems-dígitos y prueba con sondeo. En la presente investigación hemos utilizado 52 sujetos  $\times$  40 ensayos, es decir, 2.080 «palabras-señal» cuyas respuestas permitirán obtener curvas ROC.

##### Procedimiento

Las listas de palabras fueron presentadas de forma aleatoria a cada sujeto, y el orden de presentación de las listas fue balanceado. El resto de parámetros: monitor, distancia del monitor, ángulo de visión y experimentador fueron los mismos para todos los participantes.

Se aplicó el paradigma de la TDS en memoria de reconocimiento de ítems, respuestas «SÍ-NO» más escala de estimación. La respuesta «SÍ-NO» pretende obtener el registro de la respuesta en el mínimo tiempo de reacción posible, mientras que la escala de estimación pretende una valoración subjetiva sobre la fuerza del ítem en memoria: 5 muy seguro, 4 bastante seguro, 3 medianamente seguro, 2 poco seguro, 1 nada seguro. Esta escala permite construir una curva ROC con 5 puntos y 4 grados de libertad ( $5-1=4$ ).

Como hemos indicado anteriormente se eligieron 40 palabras para cada condición experimental. La mitad de las palabras, 20 para cada condición, fueron consideradas señal y otras 20 ruido. Cada palabra permanecía 500 msec en pantalla. Este incremento del tiempo de exposición, respecto a experimentos anteriores (Pelegrina y Tejeiro, 2006), pretende maximizar los efectos de la VI. Cada participante leyó las instrucciones hasta su comprensión. Además, se realizó una prueba de entrenamiento con objeto de igualar en lo posible su forma de actuación y el uso adecuado de todos los puntos de la escala. Para ello se utilizaron listas de 6 palabras (o menos si el participante entendía la tarea) diferentes a las propias del experimento.

Una vez presentada la primera lista (palabras señal) se procedió de forma inmediata (intervalo de retención inmediato) a la presentación de la segunda lista (palabras señal + ruido). Al acabar la lista señal el participante pulsaba un botón del teclado para leer las instrucciones (forma de respuesta) e iniciar la prueba de reconocimiento. Todo ello se realizó mediante el programa citado, E-prime (Schneider, Eschman y Zuccolotto, 2002) y duraba unos 30 seg. En el reconocimiento de ítems, la palabra diana permanecía en pantalla hasta que el participante pulsaba, lo más rápido posible, la tecla correspondiente a SÍ o NO (1 y 2 a la derecha del teclado, respectivamente). Sobre la precisión y rapidez de respuestas es útil ver McNicol y Steward (1980). Este tipo de escalamiento subjetivo ya fue propuesto por Torgerson (1958) y se aplica habitualmente en memoria de reconocimiento.

Tabla 1  
Palabras utilizadas en el estudio

Palabras de clase abierta		Palabras de clase cerrada			
villa	edad	lado	tuya	donde	algo
toro	falta	agua	acá	según	usted
regla	grupo	hora	mía	fuera	ahora
santa	joven	amor	ayer	otra	eso
odio	campo	tarde	abajo	cual	ello
torre	autor	forma	acaso	allí	bien
red	duda	caso	jamás	aún	así
rosa	arte	fin	cerca	tal	desde
labor	fondo	año	esta	aquí	tan
nota	aire	cosa	cuya	mucho	otro
ruido	cara	casa	lejos	antes	solo
brazo	gente	día	allá	tanto	como
carta	hijo		junto	poco	
color	calle		quizás	otro	

Análisis de datos

A los datos obtenidos se aplicó un análisis de varianza (ANOVA). La forma de la tabla de datos es equivalente a la publicada por Pelegrina y Tejeiro (2006). Se consideraron los aciertos, promediados y acumulados para cada participante.

Posteriormente, y sobre los mismos datos, se calcularon los parámetros ROC. Dichos parámetros se realizaron conjuntamente para cada condición experimental. En este sentido, se obtuvieron las curvas ROC y z-ROC, así como el índice  $d'$  y el área bajo la curva normal. Finalmente, se obtuvieron los valores de la pendiente (falsas alarmas/aciertos) y de la constante cuadrática del modelo de regresión.

Resultados

Para el efecto clase de palabras se obtuvo significación estadística en los cinco niveles de la escala de estimación (EE). Estos correspondieron a los aciertos: EE1:  $F(1, 42) = 11,29, MCE = 1,29, p = 0,03$ ; EE2:  $F(1, 42) = 7,2, MCE = 2,03, p = 0,01$ ; EE3:  $F(1, 42) = 6,18, MCE = 1,94, p = 0,001$ ; EE4:  $F(1, 42) = 12,63, MCE = 3,29, p = 0,002$ ; EE5:  $F(1, 42) = 10,34, MCE = 1,85, p = 0,02$ . Los valores promedio de las condiciones experimentales se presentan en la tabla 2.

Los cálculos de los parámetros ROC se observan en la tabla 2.

El análisis de los parámetros y curvas ROC (tabla 3) presenta una  $d'$  que crece progresivamente en el siguiente orden:  $CC < CA$ ;  $d' = 0,66 < 0,82$  ( $t_{51} = 3,94, p = 0,001$ ). Su correspondiente medida no paramétrica es consistente con este resultado: área bajo la curva ROC =  $0,67 < 0,72$  ( $t_{51} = 2,53, p = 0,014$ ). El valor negativo de las constantes cuadráticas (cc) de las curvas ROC muestran que éstas son convexas:  $cc = -0,32 < -1,80$  ( $t_{51} = 17, p < 0,001$ ). La comparación contra el valor 0 es: ( $t_{51} = 6,4, p = < 0,001$ ) y ( $t_{51} = 38,1, p < 0,001$ ), respectivamente. Vemos un crecimiento, según una ten-

dencia negativa, pero en el mismo sentido de los datos anteriores. Por el contrario, mediante las curvas z-ROC las cc son positivas y se estandariza la tendencia anterior  $0,48 > 0,34$  ( $t_{51} = 2,4, p = 0,034$ ). La comparación contra el valor 0 es: ( $t_{51} = 11,7, p < 0,001$ ) y ( $t_{51} = 7,9, p < 0,001$ ), respectivamente. Además, los resultados de la pendiente de los mismos valores de la curva z-ROC ofrece dos tipos de valores: uno de ellos, la clase cerrada, es más próximo a la unidad de pendiente,  $CC = 0,82$  que el correspondiente a la clase abierta,  $CA = 0,76$ . La diferencia entre ambos valores es significativa; ( $t_{51} = 5,47, p < 0,001$ ), al comparar la CC con la unidad: ( $t_{51} = 1,49, p = 0,142$ ), y en la CA comparada con la unidad ( $t_{51} = 26, p < 0,001$ ).

La pendiente es un indicador nítido y muy riguroso que se refiere a la presencia de umbral, o bien la varianza o asimetría de los datos. Considerando todos los resultados y si observamos la figura 2 (A, B, C y D) se asumen distribuciones con evidencia de resultados según la TDS y los modelos de umbral, así como asimetría respecto a la diagonal negativa en las curvas z-ROC.

En efecto, la figura 2A sigue un modelo de umbral mejor que 2B (curvas ROC). Por otra parte, la inflexión de la curva hacia una forma de U se observa mejor en 2C que en 2D (curvas z-ROC), pero ambas curvas z-ROC son asimétricas respecto a la diagonal negativa. Así se observa especialmente en la figura 2C con una U claramente asimétrica.

Discusión y conclusiones

Los resultados del ANOVA confirman nuestro planteamiento inicial de que el efecto de CA es diferente al efecto de CC. Por otra parte, la tendencia de las curvas ROC y z-ROC es diferente para cada condición experimental (clase de palabras). A partir de aquí inferimos que el acceso al léxico sigue un doble proceso, uno para las palabras de CA y otro para las palabras de CC. Por ello, asumimos que la varianza correspondiente a la CA (fig. 2A y C) representa efectos distintos al correspondiente de CC, más próximo a un procesamiento por familiaridad (fig. 2B y D).

Así, en la figura 2D observamos una representación del proceso de familiaridad, mediante la evidencia estimada de una z-ROC con una tendencia próxima a la linealidad (véase Glanzer, Hilford y Kim, 2009). La familiaridad es un proceso relativamente rápido que se refiere a la fuerza de un ítem en memoria. Tradicionalmente se ha modelado estadísticamente con palabras de alta frecuencia de uso mediante una distribución propia de la TDS, véase además Atkinson y Juola (1974), Jacoby y Dallas (1981) y Mandler (1980), entre otros. En general, los estímulos muy familiares, por ejemplo, el reconocimiento de rostros (v.g., Pitarque, Algarabel y Aznar, 2007), palabras muy frecuentes, o acciones muy cotidianas y comunes) se vienen ajustando al modelo clásico de la TDS. Pero

Tabla 2

Resultados experimentales en CA y CC de palabras en los cuatro momentos de la escala (el 5 es la unidad). Se representan los aciertos y las falsas alarmas en proporciones y en puntuaciones z. Son la base del ANOVA y de las curvas ROC

CA	A	0,03	0,31	0,57	0,79
	z	-1,88	-0,49	0,18	0,81
	FA	0,02	0,16	0,25	0,48
	z	-2,05	-0,99	-0,67	-0,05
CC	A	0,04	0,25	0,54	0,76
	z	-1,74	-0,67	0,10	0,71
	FA	0,03	0,09	0,34	0,52
	z	-1,88	-1,34	-0,41	0,05

Tabla 3

Parámetros e índices ROC y z-ROC para las dos clases de palabras

Palabras	Índices		ROC		z-ROC		
	$d'$	Área bajo la curva ROC	cc	$R^2$	cc	$R^2$	Pendiente
			Media		Media		Media
CA	0,82	0,72	-1,80	0,93	0,48	0,98	0,76
CC	0,66	0,67	-0,32	0,96	0,34	0,97	0,82

en palabras de CA (fig. 2C) hemos obtenido distribuciones que se identifican con procesos de *recollection* («recuerdo» en este contexto) según Yonelinas (1999) y que aquí denominamos «umbral».

Hemos de indicar que hay muchos autores, muy distantes en el tiempo, que han mantenido a lo largo de muchos años modelos continuos de respuesta *versus* umbral (v.g., Coombs, Dawes y Tversky, 1970; Egan, 1958; y recientemente Wixted, 2007). Este último autor ha propuesto un modelo de detección de señales de desigual varianza, «unequal-variance signal detection (UVSD) model», y afirma que los procesos de decisión en memoria no representan un proceso puro, sugiriendo la necesidad de investigar en diferentes vías los «múltiples» correlatos de la memoria de reconocimiento (entrecomillado nuestro).

Sobre nuestros resultados, y en relación a las teorías explicativas, podemos asumir que los estímulos muy familiares se ajustan a un modelo continuo de respuestas predicho por la TDS (Bröder y Schütz, 2009). Aunque estos autores prefieren el término *evidence strength* (p. 587) al término familiaridad, porque *fuerza* es un término con menos implicaciones teóricas. *Strength* es el término utilizado tradicionalmente en memoria mediante el paradigma de la TDS (Wickelgren y Norman, 1966).

Por otra parte, se han propuesto modelos mixtos que integran la TDS y el umbral, entendidos éstos como modelos de la TDS con desiguales varianzas (Wixted, 2007). O bien que los resultados de las curvas ROC pueden ser explicados por ambos modelos: la TDS y los modelos de umbral (Bröder y Schütz, 2009). Pero más recientemente Yonelinas y Parks (2007) proponen que los modelos deben

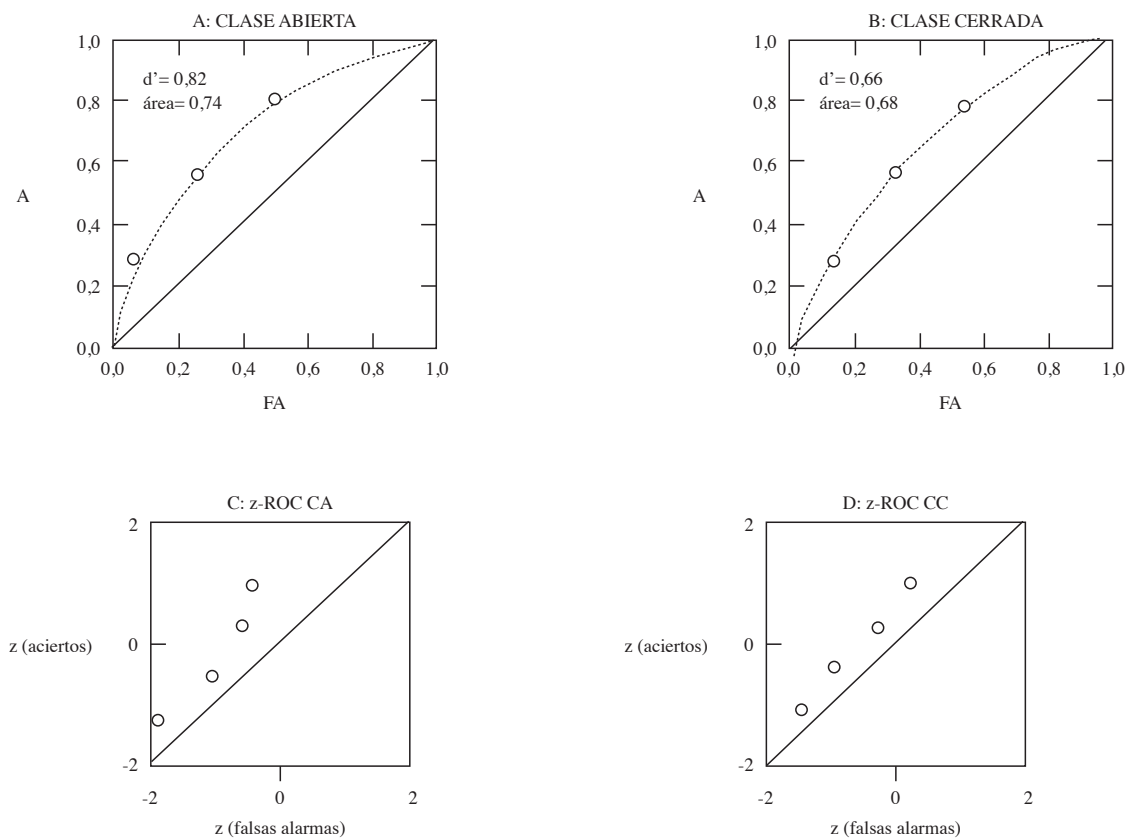
dar cuenta de los «datos existentes en las curvas ROC» empíricas (p. 198, traducción nuestra).

Otras investigaciones aportan evidencia empírica de que en el reconocimiento de ítems hay que incorporar modelos que van más allá de la familiaridad (v.g., Boldini, Russo y Avons, 2004) y en dirección a una memorización (o aprendizaje) de señales con significado (por ejemplo, palabras de CA) y ello se explica mejor desde un modelo de detección de señales de doble varianza (Jang, Wixted y Huber, 2009).

Una posible salida teórica se podría referir al hecho de que el continuo es acumulativo de evidencias en una línea determinada, que cuando llega a «un nivel de saturación» provocaría una ruptura del continuo hacia una asimetría (con diversificación de varianzas) o bien a la formación de un umbral, dependiendo todo ello de las características de los estímulos.

En cuanto a las curvas ROC se observa que representan un modelo procedimental. En efecto, vemos un paralelismo entre FA y FB (Pelegrina y Tejeiro, 2006) de palabras frente a CC y CA de palabras, respectivamente. Las curvas son similares, aunque se refieran a procesos distintos desde un modelo sustantivo: procesos más bien asociativos relacionados con la CC y la FA y procesos más bien cognitivos relacionados con la CA y la FB.

Aceptamos la TDS y las curvas ROC como modelos procedimentales aplicados a la evaluación de diferentes ítems, con los que obtenemos diferentes distribuciones, de acuerdo con Pratte, Rouder y Morey (2010). En efecto, los modelos procedimentales (que incluyen el recuerdo y la familiaridad en este contexto (Mic-



**Figura 2.** A, B, C y D. A: curva ROC empírica obtenida mediante EE de 5 puntos correspondiente a palabras de CA, calculada a partir de las respuestas de todos los participantes en la correspondiente condición experimental. B: el mismo resultado correspondiente a la CC de palabras. C: curva z-ROC estimada a partir de la figura A. D: curva z-ROC estimada a partir de la figura B. Se observa una diferente asimetría respecto a la diagonal negativa en C y en D

kes, Johnson y Wixted, 2010) son más plausibles en orden a su mejor integración que si se partiese exclusivamente de un modelo sustantivo (perceptual o sensorial) de la TDS.

Hemos de terminar añadiendo las múltiples aplicaciones de los modelos propuestos para evaluar variables de memoria y lenguaje especialmente en ámbitos educativos y clínicos. Como ejemplo de

partida proponemos la lectura del monográfico de la revista *Psicothema* en 1999, véase Baqués y Sáiz (1999) en el aprendizaje de la lectura, y Peraíta, Galeote y González (1999) en pacientes con Alzheimer. Todo ello con posibilidades de medida aplicadas según el modelo propuesto en esta investigación y en la línea de Rattcliff, Sheu y Gronlund (1992).

## Referencias

- Alameda, J.R., y Cuetos, F. (1995). *Diccionario de frecuencias de las unidades lingüísticas del castellano*. Universidad de Oviedo. Departamento de Psicología
- Atkinson, R.C., y Juola, J.F. (1974). Search and decision processes in recognition memory. En D.H. Krantz, R.C. Atkinson, R.D. Luce y P. Suppes (Eds.), *Contemporary developments in mathematical psychology: Vol. I. Learning, memory & thinking* (pp. 242-293). San Francisco: Freeman.
- Baddeley, A.D. (1983). *Psicología de la memoria*. Madrid: Debate.
- Baqués, J., y Sáiz, D. (1999). Medidas simples y compuestas de memoria de trabajo y su relación con el aprendizaje de la lectura. *Psicothema*, 11(4), 737-745.
- Biassou, N., Obler, L.K., Nespoulous, J.L., Dordain, M., y Harris, K.S. (1997). Dual processing of open-and closes-class words. *Brain and Language*, 57, 360-373.
- Blackwell, H.R. (1963). Neural theories of simple visual discrimination. *Journal of the Optical Society of America*, 53, 129-160.
- Boldini, A., Russo, R., y Avons, S.E. (2004). One process is not enough! A speed-accuracy trade-off study of recognition memory. *Psychonomic Bulletin & Review*, 11, 353-361.
- Bröder, A., y Schütz, J. (2009). Recognition ROCs are curvilinear-or are they? On premature arguments against the two- high-threshold model recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 35(3), 587-606.
- Coombs, C.H., Dawes, R.M., y Tversky, A. (1970). *Mathematical Psychology*. New Jersey: Prentice-Hall.
- Egan, J.P. (1958). *Recognition memory and operating characteristic*. Technical Note AFCRE-IN-58-51.
- Egan, J.P. (1975). *Signal detection theory and ROC analysis*. New York: Academic Press.
- Glanzer, M., Hilford, A., y Kim, K. (2004). Six regularities of source recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 30(6), 1176-1195.
- Gorman, A.M. (1961). Recognition memory for nouns as a function of abstractness and frequency. *Learning, Memory and Cognition*, 25, 500-513.
- Green, D.M., y Swets, J.A. (1966). *Signal detection and psychophysics*. New York: John Wiley.
- Jang, Y., Wixted, J.T., y Huber, D.E. (2009). Testing signal-detection models of yes-no and two alternative forced-choice recognition memory. *Journal of Experimental Psychology: General*, 138(2), 291-306.
- Jacoby, L.L., y Dallas, M. (1981). On the relationship between autobiographical memory and perceptual learning. *Journal of Experimental Psychology: General*, 2(110), 139-154.
- Johnson, V.E. (1999). On bayesian analysis of multirater ordinal data: An application to automated essay grading. *Journal of the American Statistical Association*, 91(433), 42-51.
- Juila, A., y Chang-Rodríguez, E. (1964). *Frequency dictionary of Spanish words*. Mouton and Co.: La Haya.
- MacMillan, N.A., y Creelman, C.D. (1990). Response bias: Characteristics of detection theory, threshold theory and «non parametric» indexes. *Psychological Bulletin*, 107, 401-413.
- Mandler, G. (1980). Recognizing: The judgment of previous occurrence. *Psychological Review*, 87, 252-271.
- McNicol, D., y Steward, G.W. (1980). Reaction time and the study of memory. En A.T. Welford (Ed.), *Reaction times* (pp. 253-307). New York: Academic Press.
- Mickes, L., Johnson, E.M., y Wixted, J.T. (2010). Continuous recollection versus unitized familiarity in associative recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 36(4), 843-863.
- Münte, T.F., Wieringa, B.M., Weyerts, H., Szentkuti, A., Matzke, M., y Johannes, S. (2001). Differences in brain potentials to open and closed class words: Class an frequency effects. *Neuropsychology*, 39, 91-102.
- Pelegrina, M., y Tejeiro, R. (2006). Parámetros ROC y z-ROC en memoria de palabras: efectos experimentales y preexperimentales. *Psicothema*, 18(1), 160-164.
- Peraíta, H., Galeote, M.A., y González, M.J. (1999). Deterioro de la memoria semántica en pacientes de Alzheimer. *Psicothema*, 11(4), 917-937.
- Pitarque, A., Algarabel, S., y Aznar-Casanova, J.A. (2007). Familiaridad y recuerdo en el reconocimiento de rostros ficticios: implicaciones de los modelos de reconocimiento. *Psicothema*, 19, 565-571.
- Pratte, M.S., y Rouder, J.N. (2010). Separating mnemonic process from participant and item effects in the assessment of ROC asymmetries. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 36(1), 224-232.
- Rattcliff, R., Sheu, C., y Gronlund, S.D. (1992). Testing global memory models using ROC curves. *Psychological Review*, 99, 518-535.
- Schneider, W., Eschman, A., y Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh: Psychology Software Tools Inc.
- Torgerson, W.S. (1958). *Theory and methods of scaling*. New York: Wiley.
- Yonelinas, A.P. (1999). The contribution of recollection and familiarity to recognition and source-memory judgments: A formal dual-process model and an analysis of receiver operating characteristics. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 25(6), 1415-1434.
- Yonelinas, A.P., y Parks, C.M. (2007). Receiver operating characteristics in recognition memory: A review. *Psychological Bulletin*, 133(5), 800-832.
- Wickelgren, W.A., y Norman, D.A. (1966). Strength models and serial position in short-term memory. *Journal of Mathematical Psychology*, 3, 316-347.
- Wixted, J.T. (2007). Dual-process theory and signal-detection theory of recognition memory. *Psychological Review*, 114(1), 152-176.