

Predicciones bajo hipótesis de distribuciones gaussianas y laplacianas en un modelo CTVA-2D

Carmen Santisteban Requena y Jesús M^a Alvarado Izquierdo
Universidad Complutense de Madrid

La teoría CODE de Atención Visual que presenta Logan en 1996 (Psychological Review, 103, 603-649) considera que la posición que ocupan los elementos en una presentación visual, en lugar de representarse por puntos en el espacio, se representan por distribuciones de probabilidad. Logan desarrolla un modelo utilizando distribuciones Laplacianas unidimensionales. Este modelo se ha extendido a dos dimensiones generando el modelo CTVA-2D (Alvarado, Santalla y Santisteban, 1999, Acta Psychologica, 103, 239-255). En este trabajo se comparan los resultados bajo los supuestos de distribuciones gaussianas y de laplacianas en la citada extensión bidimensional del modelo.

Predictions in a CTVA-2D model under normal and Laplace's distributions hypotheses. The CODE theory of Visual Attention introduced by Logan in 1996 (Psychological Review, 103, 603-649) assumes that the items location is not represented by points in the space whereas it is represented by one-dimensional Laplace distributions. This model has been extended to two dimensions by the CTVA-2D model (Alvarado, Santalla & Santisteban, 1999, Acta Psychologica, 103, 239-255). The present work compares the results obtained assuming Laplace and normal distributions in the two-dimensional model.

La teoría CODE (Contour Detector) de agrupamiento perceptivo por proximidad propuesta inicialmente por van Oeffelen y Vos (1982, 1983) contiene como uno de sus aspectos claves la suposición de que la localización de cada uno de los elementos no es puntual sino que son distribuciones en el espacio. Así, la probabilidad de detección de cada elemento en el espacio se puede representar por su propia distribución de probabilidad y desde una perspectiva de procesamiento del objeto al sujeto («bottom-up»), la suma de las distribuciones de los diferentes elementos que se presentan, generan la superficie CODE.

Las distribuciones, en principio, pueden ser cualesquiera si se considera que son simétricas, aunque sea de forma grosera, y apuntadas en el centro. Inicialmente la teoría CODE, supuso que las distribuciones probabilísticas eran normales. Posteriormente, Compton y Logan (1993) muestran que presentando patrones de puntos, el uso de distribuciones unidimensionales de Laplace en la superficie CODE, dan mejores predicciones acerca de los juicios de agrupamiento que hacen los sujetos que si se hipotetiza que las distribuciones son normales.

La teoría de la atención selectiva formulada por Logan (1996) partiendo de la CODE y de la Teoría de Atención Visual (TVA) de Bundesen (1990) también considera a priori distribuciones laplacianas para representar la distribución probabilística de la localización de cada uno de los puntos. El modelo que formalmente recoge esta teoría CODE de atención visual es el modelo CTVA

(CODE Theory of Visual Attention) que formula Logan (Logan, 1996) para una dimensión. La extensión a dos dimensiones, generando la CTVA-2D, también se ha realizado bajo los supuestos de que las distribuciones probabilísticas de la identificación en cada punto son laplacianas (Alvarado, Santalla y Santisteban, 1999).

El objetivo del presente trabajo es comparar los resultados bajo los supuestos de distribuciones gaussianas y de laplacianas en la extensión bidimensional del modelo CTVA.

Formulación

La superficie CODE proporciona información espacial, sobre las distribuciones probabilísticas de las ubicaciones de los elementos o ítems individuales y de las distribuciones de los objetos, o grupos de objetos, que se pueden formar al procesarse conjuntamente elementos próximos entre sí.

La representación de las distribuciones probabilísticas de tres elementos se ilustra en la Figura 1. Los puntos x , y , z representan la localización en la presentación de tres elementos en la dimensión S . Las funciones $f_x(s)$, $f_y(s)$ y $f_z(s)$ son las respectivas funciones de densidad. La superficie CODE, que viene definida por la suma de las distribuciones correspondientes a los elementos individuales, es la encerrada por la curva $F(s)$ y el eje horizontal.

Fijado un umbral $f_i(s) = K$ la percepción de objetos diferenciados o grupos de objetos dependerá de ese valor K y de sus puntos de corte con $F(s)$. En la Figura 1 se percibirían dos grupos, uno lo forman los objetos (y , z) que están próximos, y otro formado por el elemento x . Los diferentes elementos se identifican muestreando sus características con una probabilidad igual al área encerrada por la curva $F(s)$ y la función $f_i(s) = K$, región que en la literatura se denomina «feature catch». También se puede observar en la Figura 1 que la probabilidad o «feature catch» de muestrear las ca-

racterísticas del grupo (y, z) es alta, puesto que incluye la mayor parte de la densidad de probabilidad de las distribuciones de estos dos elementos. Una parte muy pequeña de la distribución asociada al elemento x, no agrupado con los otros dos elementos, cae dentro de la «feature catch», lo que podría provocar que alguna característica de este elemento se muestree junto a las características de los dos elementos agrupados, con una probabilidad proporcional al área de la distribución que cae en éste área delimitada.

La superficie CODE en una presentación bi-dimensional de N elementos (CODE-2D), suma de las distribuciones de cada uno de ellos estará representada por las ecuaciones (1) o (2) dependiendo de que las distribuciones probabilísticas básicas sean:

a) Distribuciones laplacianas:

$$h(x,y) = \sum_{i=1}^N h_i(x, y) = \sum_{i=1}^N \frac{\exp\left[-2\left(\frac{r_x^2 + r_y^2}{\sigma_x^2 \sigma_y^2 - \rho_{xy}^2}\right)\right]^{1/2}}{\left[2(\sigma_x^2 \sigma_y^2 - \rho_{xy}^2)\right]^{1/2}} \quad (1)$$

b) Distribuciones gaussianas:

$$h(x,y) = \sum_{i=1}^N h_i(x, y) = \sum_{i=1}^N \frac{\exp\left\{-\frac{1}{2(1-\rho_{xy}^2)}\left[\left(\frac{r_x}{\sigma_x}\right)^2 + \left(\frac{r_y}{\sigma_y}\right)^2 - 2\rho_{xy}\left(\frac{r_x}{\sigma_x}\right)\left(\frac{r_y}{\sigma_y}\right)\right]\right\}}{2\pi\sigma_x\sigma_y\left[1-\rho_{xy}^2\right]^{1/2}} \quad (2)$$

siendo μ_x, μ_y y σ_x, σ_y respectivamente las medias y las desviaciones típicas de las distribuciones marginales de cada elemento i en la presentación. Los vectores $r_x = (x - \mu_x)$ y $r_y = (y - \mu_y)$ son la coordenadas espaciales del elemento i . En este trabajo se asume que las desviaciones típicas marginales en ambas dimensiones son iguales ($\sigma_x = \sigma_y$) y que la covarianza es cero $\rho_{xy} = 0$.

En la Figura 2, se presenta una superficie CODE bidimensional de tres elementos para distribuciones gaussianas y laplacianas. En este caso se presenta un objetivo flanqueado por dos elementos distractores, que es la disposición clásica de los elementos en el paradigma de Eriksen (Eriksen y Eriksen, 1972).

Construida la superficie CODE y calculada la «feature catch», se puede hacer uso de las fórmulas propuestas por la TVA de Bundesen (1990) para el cálculo de la tasa de categorización. La probabilidad de respuesta correcta y la predicción de los tiempos de

reacción se realiza a través de los cálculos de esa tasa. De acuerdo con la TVA, se define la tasa de categorización, $v(x,i)$ de un elemento x perteneciente a la categoría i como:

$$v(x,i) = C_x \eta(x,i) \beta_i \frac{w_x}{\sum_{z \in S} w_z} = C_x \eta(x,i) \beta_i \frac{\sum_{j \in R} \eta(x, j) \pi_j C_x}{\sum_{z \in S, j \in R} \sum \eta(z, j) \pi_j C_z}$$

siendo C_x la *feature catch* del elemento x ; $\eta(x,i)$ la fuerza de la evidencia sensorial de que el elemento x pertenezca a la categoría i ; β_i la predisposición perceptiva asociada a la categoría i y siendo w_x y w_z los respectivos pesos atencionales de los elementos x, z , pesos que se derivan de los valores π_j , que representan la importancia de atender a los elementos de cada categoría j .

La predicción de los tiempos de reacción de la respuesta correcta, TR_c se obtienen mediante la expresión:

$$TR_c = \left\{ \sum_{j=0}^{K_p-1} \binom{K_p + j - 1}{j} \left(\frac{v(x, p)}{v(x, p) + v(x, q)} \right)^{K_p} \right.$$

$$\left. \times \left(\frac{v(x, q)}{v(x, p) + v(x, q)} \right)^j \binom{K_p + j}{j} \left[\frac{1}{P(R_c)} \right] \right\} + b$$

donde b es una constante y $P(R_c)$ representa la probabilidad de respuesta correcta cuya expresión sería:

La formulación anteriormente expuesta se ha extendido al modelo CTVA-2D (Alvarado, Santalla y Santisteban, 1999) y se ha aplicado a los datos experimentales obtenidos con la extensión del paradigma de Eriksen (Eriksen y Eriksen, 1974) realizada por los mismos autores (Alvarado, Santalla y Santisteban, 1998).

Para comparar las predicciones bajo las hipótesis de normalidad con respecto a los obtenidos con distribuciones de Laplace, que es el objeto del presente estudio, se ha utilizado el método que se explicita a continuación.

Método

Sujetos

Cinco hombres y tres mujeres, entre 22 y 34 años, licenciados en Psicología por la Universidad Complutense de Madrid, participaron como voluntarios. Todos eran diestros y tenían visión normal o corregida.

Aparatos y estímulos

Para presentar los estímulos y recoger las respuestas de los sujetos se utilizó un ordenador personal. Los estímulos presentados y las respuestas recogidas fueron controlados por un programa en C++ (Alvarado, 1997). Los sujetos dieron sus respuestas mediante el teclado del ordenador.

La tarea consistía en identificar la letra objetivo («p», «b», «q» o «d») que aparecía en el centro de la pantalla del monitor, rodea-

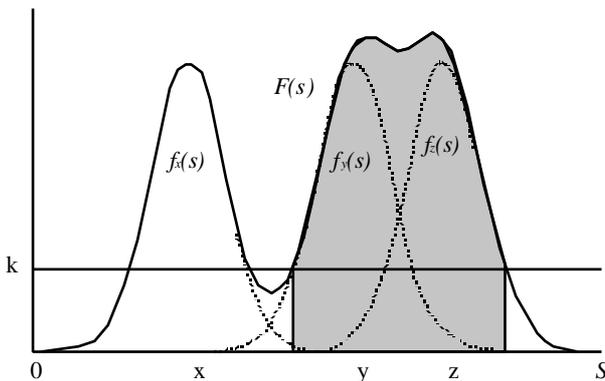


Figura 1. Bajo supuestos de normalidad: funciones de densidad de probabilidad $f_i(s)$, superficie CODE $\int_0^s F(s) ds$ de los elementos x, y, z , y «feature catch» (sombreado) de los elementos y, z .

da por dos, cuatro u ocho distractores. Los sujetos debían pulsar el cursor derecho si la letra objetivo era una «p» o una «b», y el cursor izquierdo si era una «q» o una «d». Tanto el objetivo como los distractores aparecían simultáneamente. Todas las letras eran minúsculas de color blanco (luminancia: 116 cd/m²) presentadas sobre un fondo negro (luminancia: 1 cd/m²).

Se consideraron tres variables independientes:

(1) La compatibilidad entre objetivo y distractores. La variable compatibilidad podía adoptar una de las siguientes modalidades: (i) Respuesta compatible, cuando el objetivo es «p» o «b» los distractores deben ser «p» o «b», cuando el objetivo es «q» o «d» los distractores deben ser «q» o «d» (ii) Respuesta incompatible, cuando el objetivo es «p» o «b» los distractores deben ser «q» o «d», cuando el objetivo es «q» o «d» los distractores deben ser «p» o «b» y (iii) condición neutra, cuando los distractores no eran posibles letras objetivo, se utilizaron las letras «x» y «z»).

(2) El número y ubicación de los distractores, o configuración: La variable configuración podía adoptar uno de las tres modalidades que se muestran en la Figura 3 según se presentaran dos, cuatro u ocho distractores.

(3) La distancia entre el objetivo y los distractores considerando las distancias 0.97°, 1.95°, 2.92° y 3.89° en grados de ángulo visual.

Diseño

Se utilizó un diseño factorial de medidas repetidas. Se experimentó con los 36 tipos diferentes de presentaciones obtenidas por combinación de las distintas modalidades de las tres variables independientes consideradas (3 x 3 x 4). Cada una de ellas se presentó 20 veces a cada sujeto en una sesión experimental. El orden de presentación de los 720 (20 x 36) ensayos fue aleatorio. En cada ensayo se registró el tiempo de reacción y si la respuesta era correcta o incorrecta.

Procedimiento

Los sujetos realizaron la tarea en una cámara insonorizada, reposando su cabeza en un barbuquejo situado a 60 cm de la pantalla del monitor. La hora del día, iluminación y otras condiciones ambientales se mantuvieron constantes.

Cada ensayo comenzaba con un punto de fijación (+) que permanecía durante 1000 ms en el centro del monitor, sustituyéndose posteriormente por el objetivo. La presentación permanecía hasta que el sujeto respondía y pulsaba el espaciador dando comienzo el siguiente ensayo. Cuando los sujetos cometían un error se les daba un feedback sonoro (un tono de 500 Hz). Los sujetos comenzaban con una sesión de práctica que consistía en la realización de 90 ensayos aleatoriamente seleccionados entre los 720 considerados. Las respuestas dadas en esta sesión de práctica no se incluyeron en los análisis.

Los sujetos fueron instruidos para responder solo al objetivo que aparecía en el centro de la pantalla, respondiendo tan rápido como les fuera posible y evitando cometer errores.

Resultados

Se computó separadamente las medias de los tiempos de reacción (TRs) y de los errores en cada una de las 36 presentaciones diferentes posibles para cada participante, realizándose un ANOVA de medidas repetidas intra-sujeto, considerando los siguientes factores: a) compatibilidad de los distractores con el objetivo b) configuración y c) distancia entre el objetivo y los distractores. Únicamente los TRs menores de 1000 ms en los casos de respuesta correcta se incluyeron en el análisis, descartándose por ese motivo el 1.81% de los ensayos.

Se encontró un efecto principal estadísticamente significativo de dos de los tres factores considerados: compatibilidad de los distractores ($F_{(2,14)} = 100.84, p < 0.01$), y distancia ($F_{(3,21)} = 56.82, p < 0.01$). La interacción entre distancia y compatibilidad también fue estadísticamente significativa ($F_{(6,42)} = 9.23, p < 0.01$), así como la interacción entre configuración y compatibilidad ($F_{(4,28)} = 4.31, p < 0.01$).

Estos resultados con efectos significativos se presentan gráficamente (Figuras 4-7). En cada una de las figuras se presentan diferentes gráficos que se corresponden con los resultados experimentales y con las predicciones de los TRs del modelo CTVA-2D con distribuciones laplacianas y con distribuciones gaussianas.

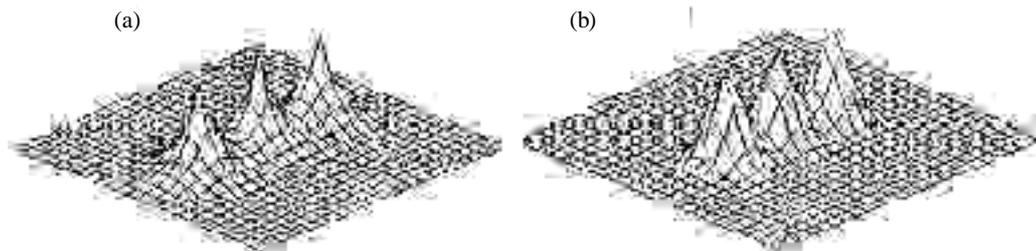


Figura 2. Superficie CODE bidimensional de tres elementos: (a) distribuciones básicas laplacianas y (b) distribuciones básicas gaussianas

(a)	(b)	(c)
b p b	q q p q q	z z z z p z z z z

Figura 3. Ejemplos de tres configuraciones: (a) respuesta-compatible, (b) respuesta-incompatible, y (c) condición neutra

Conclusiones

El uso de distribuciones gaussianas y laplacianas para conformar la superficie CODE-2D, permite hacer buenas predicciones de los TRs en la tarea propuesta a los sujetos. No obstante, el uso de distribuciones laplacianas se muestra más favorable en casi todos los casos analizados. En concreto para las tres variables independientes que se explicitan en este trabajo, se concluye que:

1. Cuando se toma en consideración la variable compatibilidad, ambos modelos en 2D se comportan de forma similar dando buenas predicciones.

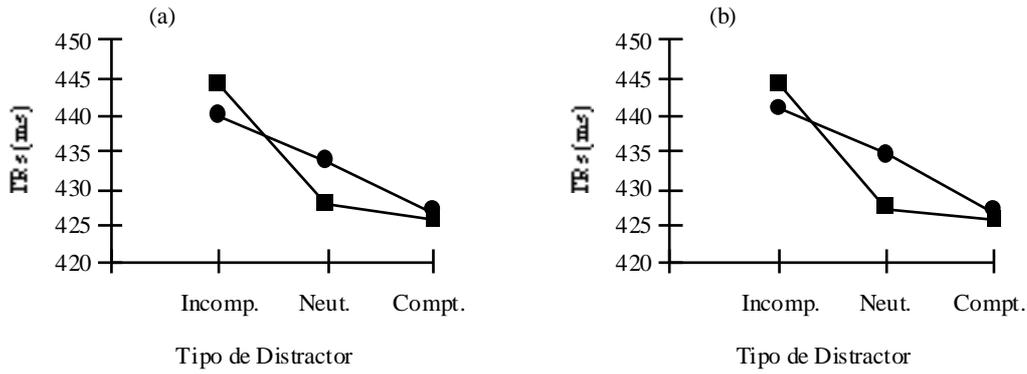


Figura 4. TRs predichos (●) y observados (■) en función del tipo de distractor para: (a) distribuciones laplacianas y (b) distribuciones gaussianas

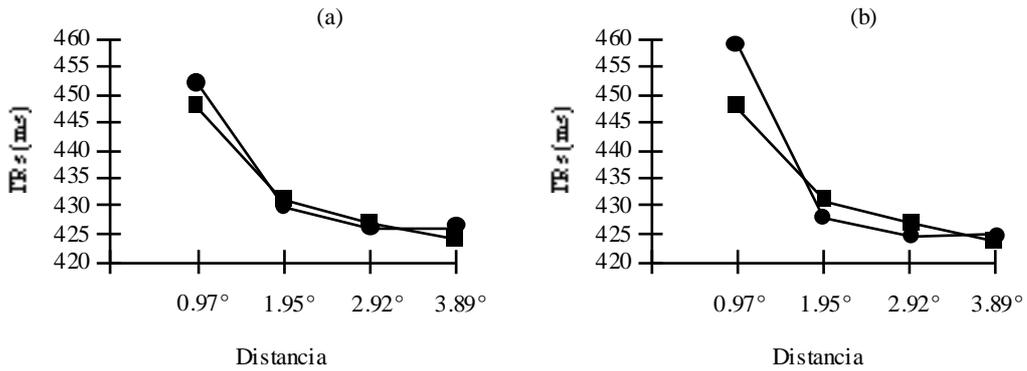


Figura 5. TRs predichos (●) y observados (■) en función de la distancia entre el objetivo y los distractores para: (a) distribuciones laplacianas y (b) distribuciones gaussianas

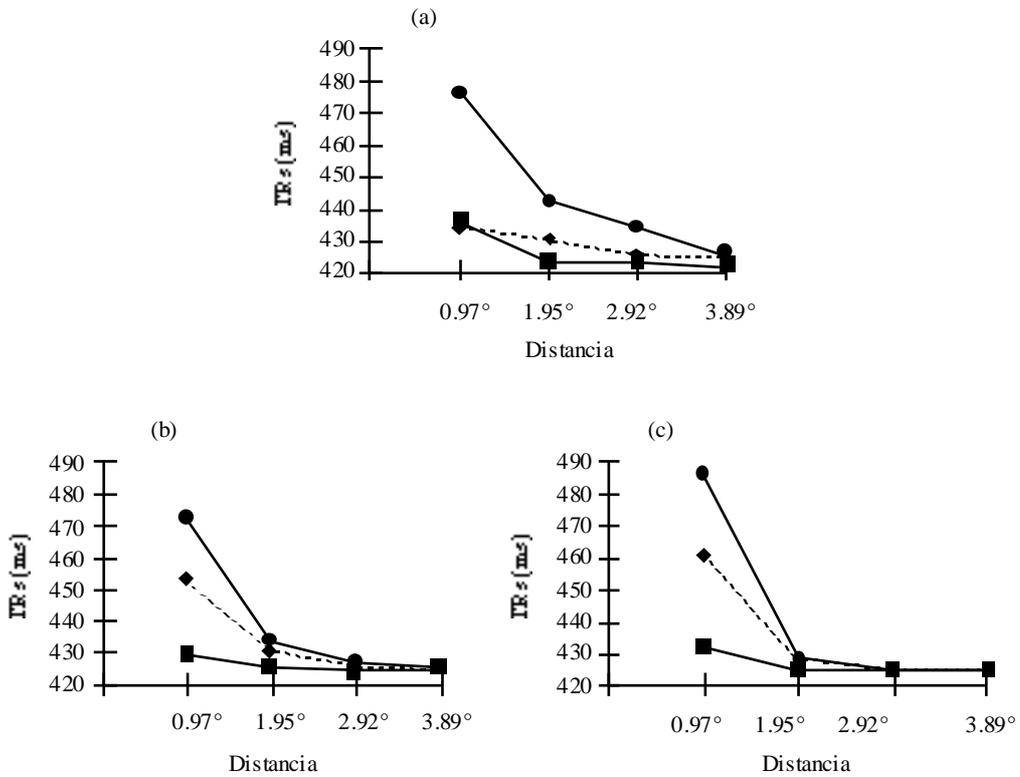


Figura 6. TRs en función de la distancia y el tipo de distractor: respuesta-incompatible (●), respuesta-compatible (■) y neutra (◆). En (a) se presentan los TRs observados, en (b) los que se predicen con distribuciones laplacianas y en (c) los que se predicen con distribuciones gaussianas

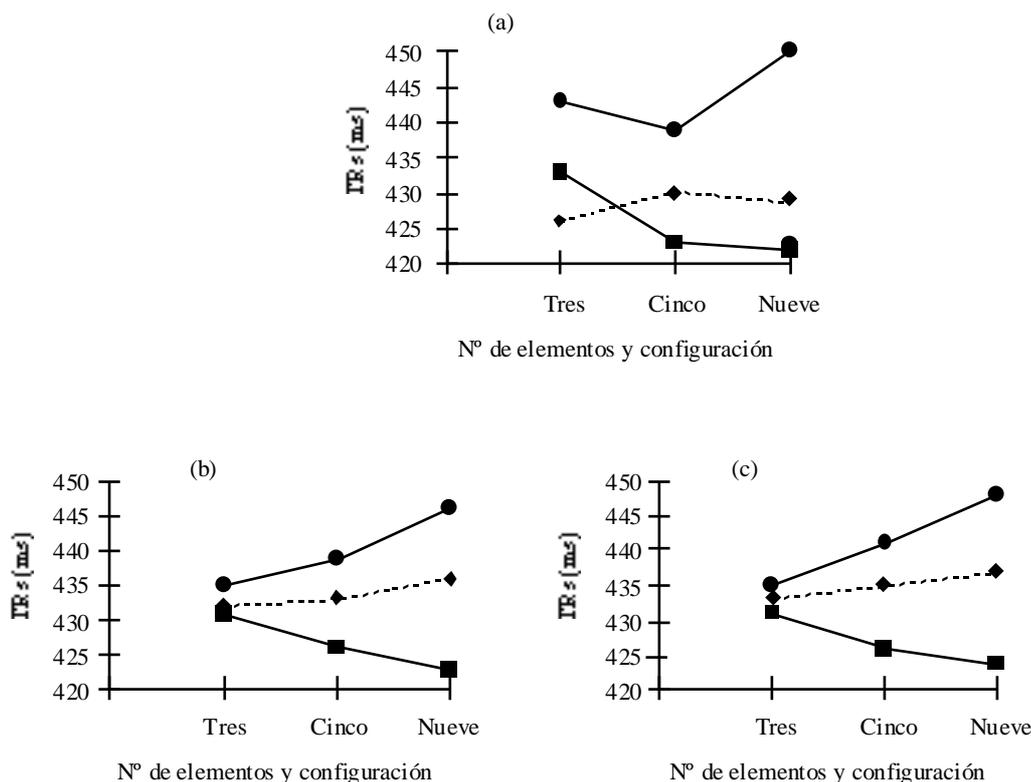


Figura 7. TRs en función de la configuración (número de elementos) y el tipo de distractor: respuesta-incompatible (●), respuesta-compatible (■) y neutra (◆). En (a) se presentan los TRs observados, en (b) los que se predicen con distribuciones laplacianas y en (c) los que se predicen con distribuciones gaussianas

2. Cuando se considera aisladamente la variable distancia, las predicciones son mucho mejores al utilizar distribuciones laplacianas. La distribuciones gaussianas tienden a sobrestimar el efecto del agrupamiento en la distancia más pequeña (0.97°) y a subestimar este efecto en las mas grande de las consideradas.

3. En cuanto a la interacción distancia-compatibilidad, se observa que las predicciones con el modelo gaussiano (2D) difieren de los datos experimentales en mayor medida que el modelo laplaciano (2D) debido fundamentalmente a que con distribuciones básicas gaussianas se produce una sobreestimación de los TRs en respuesta-compatible.

4. En la interacción distancia-compatibilidad, ninguno de los dos modelos parece explicar el comportamiento de los distractores neutros, que parecen comportarse de forma muy similar a los estímulos compatibles.

5. En cuanto a la interacción configuración-compatibilidad, ambos modelos predicen una reducción de los TRs al aumentar el número de elementos en respuesta compatible y un aumento de los TRs cuando aumenta el número de distractores incompatibles, lo que en líneas generales confirman nuestros datos experimentales, aunque con algunas excepciones (Alvarado, Santalla y Santisteban, 1999).

Agradecimientos

Este trabajo está parcialmente financiado por el Ministerio de Educación y Cultura, proyecto BIO97-0543 y por la Universidad Complutense de Madrid proyecto PR156/97-7193.

Referencias

- Alvarado, J.M^a. (1997). Programa para la evaluación de la competición entre respuestas (PECER 2.0). Documento interno de la Unidad de Resonancia Magnética Nuclear. Universidad Complutense de Madrid.
- Alvarado, J.M^a., Santalla, Z., y Santisteban, C., (1998). Efecto de la segregación sobre el procesamiento de la estimulación visual. *Psicológica*, 19, 87-105.
- Alvarado, J.M^a., Santalla, Z. y Santisteban, C. (1999). An evaluation of the CODE Theory of Visual Attention extended to two dimensions. *Acta Psychologica*, 103, 239-255.
- Bundesden, C., (1990). A theory of visual attention. *Psychological Review*, 97, 523-547.
- Compton, B.J., & Logan, G.D., (1993). Evaluating a computational model of perceptual grouping by proximity. *Perception & Psychophysics*, 53, 403-421.
- Eriksen, B.A., & Eriksen, C.W., (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16, 143-149.
- Logan, G.D., (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review*, 103, 603-649.
- van Oeffelen, M.P., y Vos, P.G., (1982). Configurational effects on the enumeration of dots: Counting by groups. *Memory & Cognition*, 10, 396-404.
- van Oeffelen, M.P., y Vos, P.G. (1983). An algorithm for pattern description on the level of relative proximity. *Pattern Recognition*, 16, 341-348.