

Fiabilidad y aplicaciones de una prueba de evaluación de la comprensión gramatical

Ramón López-Higes, M. Teresa Martín-Aragoneses, Susana Rubio-Valdehita

Introducción. La prueba de comprensión de oraciones de la batería ECCO (exploración cognitiva de la comprensión de oraciones) permite evaluar a niños y adultos con alteraciones de la comprensión gramatical a través de un procedimiento de verificación simple que demanda pocos recursos de memoria.

Objetivo. Presentar un análisis psicométrico de la prueba de comprensión de oraciones basado en la teoría clásica de los test y en la teoría de la respuesta al ítem, así como algunos estudios que muestran su utilidad para evaluar la comprensión gramatical en distintas poblaciones.

Sujetos y métodos. La prueba se aplicó a 2.238 sujetos de edades comprendidas entre los 6 y los 85 años.

Resultados. El análisis realizado muestra que la prueba tiene una elevada fiabilidad global o a través de los grupos de edad considerados, incluye elementos especialmente discriminativos o informativos (distractores léxicos y sintácticos), y está sesgada hacia niveles medios-bajos del rasgo.

Conclusiones. A la vista de los resultados, la prueba de comprensión de oraciones de la batería ECCO proporciona una medida fiable de la comprensión gramatical a través de una amplia variedad de construcciones, y tiene una gran utilidad potencial en el ámbito clínico y de la investigación.

Palabras clave. Comprensión gramatical. Distractor léxico. Distractor sintáctico. Evaluación del lenguaje escrito. Fiabilidad. Teoría de la respuesta al ítem. Trastornos de la comprensión.

Departamento de Psicología Básica II: Procesos Cognitivos (R. López-Higes, M.T. Martín-Aragoneses); Departamento de Personalidad, Evaluación y Tratamientos Psicológicos II: Psicología Diferencial y del Trabajo (S. Rubio-Valdehita); Universidad Complutense de Madrid. Madrid, España.

Correspondencia:

Dr. Ramón López Higes. Departamento de Psicología Básica II: Procesos Cognitivos. Facultad de Psicología. Universidad Complutense de Madrid. Campus de Somosaguas, s/n. E-28223 Pozuelo de Alarcón (Madrid).

Fax:

+34 913 943 189.

E-mail:

rlopeza@psi.ucm.es

Aceptado tras revisión externa:

06.11.09.

Cómo citar este artículo:

López-Higes R, Martín-Aragoneses MT, Rubio-Valdehita S. Fiabilidad y aplicaciones de una prueba de evaluación de la comprensión gramatical. *Rev Neurol* 2010; 50: 395-402.

© 2010 Revista de Neurología

Introducción

Las principales baterías de evaluación neuropsicológica o los test de lenguaje de carácter cognitivo incluyen subpruebas que evalúan la comprensión del lenguaje escrito. En todas ellas, las pruebas que exploran el acceso al léxico o la activación del significado de las palabras están especialmente representadas. El ejemplo más claro de ello es el test de evaluación del procesamiento lingüístico en la afasia (EPLA) [1].

Sin embargo, la comprensión de oraciones no se evalúa con la misma exhaustividad a pesar de su importancia. Predominan las pruebas que exploran la comprensión de nociones semánticas (ejecución de órdenes, comprensión de nociones espaciales, etc.), mientras que la comprensión gramatical, es decir, la asignación de los roles temáticos a partir de las claves sintácticas de las oraciones, sólo se evalúa de forma limitada en algunas pruebas, como el EPLA o el *Everyday Life Activities* [2] en el caso de adultos.

En pruebas dirigidas a niños o adolescentes se evalúa de forma más sistemática, como en el test de sintaxis de Aguado (TSA) [3], en la adaptación española del *Test for Reception of Grammar* (TROG)

[4,5], que explora la comprensión gramatical en la modalidad auditiva, o en la batería de lenguaje objetiva y criterial (BLOC) [6], que posee un módulo de sintaxis.

Una breve revisión de trabajos recientes puede servir para establecer la pertinencia y relevancia de este tema. Distintos estudios han aportado evidencias que respaldan la hipótesis de un declive en la habilidad de procesamiento sintáctico asociada a la edad, que se manifiesta en un bajo nivel de desempeño de personas mayores en tareas que requieren comprensión de oraciones complejas [7,8].

Mientras que los pacientes que presentan comprensión asintáctica [9] muestran dificultades con las oraciones que contienen dependencias de larga distancia entre constituyentes, los afásicos de conducción tienen dificultades para la reactivación fonológica del material presentado anteriormente a la hora de resolver la ambigüedad sintáctica local [10].

Por otro lado, los pacientes afásicos con mayores dificultades en la comprensión de enunciados tienen también problemas para construir la estructura sintáctica en el curso del procesamiento [11]. Sin embargo, los pacientes con demencia tipo Alzheimer con bajos resultados en la comprensión de

enunciados complejos (en tareas *'off line'*) no muestran dificultades a la hora de construir la estructura sintáctica en el curso de la comprensión [12]. Las dificultades de los pacientes con demencia tipo Alzheimer se relacionarían con el uso de la información extraída de los enunciados para realizar otros procesos postinterpretativos (juicios, etc.), que se ven afectados especialmente por el número de proposiciones contenidas en las oraciones [13,14].

Otro ejemplo sería el de las dificultades de los pacientes con enfermedad de Parkinson ante oraciones complejas, que se relacionarían con una disminución de los recursos destinados a la inhibición de conductas automatizadas o a problemas para mantener y manipular información verbal a corto plazo [15].

También se han observado estas dificultades en niños que presentan un trastorno específico del lenguaje [16], por citar alguna categoría diagnóstica en el ámbito de los trastornos del desarrollo.

En todos los casos, las dificultades en la comprensión de oraciones parecen depender de una disminución de los recursos de procesamiento disponibles en comparación con sujetos de las mismas características (edad, nivel educativo, etc.) [17,18]. Factores como la complejidad sintáctica o el número de proposiciones, entre otros, pueden poner de manifiesto una disminución anormal de los recursos dedicados al procesamiento en tareas que exploran el curso de la comprensión, y en otras que evalúan el uso de esta información.

Una prueba que explore de forma sistemática la comprensión de oraciones debería servir para determinar si un paciente puede hacer uso de las claves sintácticas con el fin de establecer el significado en una amplia variedad de construcciones, y para comprobar si las dificultades de un paciente tienen un origen léxico o si son el resultado de su baja capacidad de memoria operativa.

La batería ECCO (exploración cognitiva de la comprensión de oraciones) es una prueba aplicada mediante ordenador que explora la comprensión de oraciones, e incluye dos pruebas complementarias, una de vocabulario (conocimiento del léxico) y otra de capacidad de la memoria operativa verbal (memoria operativa).

La prueba de comprensión de oraciones (CO) permite explorar la asignación de los roles temáticos a los constituyentes [19] mediante un procedimiento que demanda pocos recursos de memoria: la verificación simple de pares de oraciones y dibujos [20]. En ella se evalúan 17 categorías de estructuras sintácticas. En cada categoría aparecen tres tipos de ítems: congruentes, distractores léxicos (cambia la acción/predicado o uno de los argumentos en el di-

bujos), y distractores sintácticos (los roles temáticos están invertidos en el dibujo). De esta manera, puede determinarse si un sujeto tiene una dificultad específica con un tipo de construcciones, ya que responderá 'verdadero' ante los distractores sintácticos. Alternativamente, si los errores se concentran en los distractores léxicos, habría que sospechar que las dificultades del sujeto tendrían una base léxica.

La prueba complementaria de memoria operativa (MO), semejante al *Reading Span Test* [21], incluye series con un número creciente de oraciones y otras series decrecientes. Las oraciones empleadas varían entre las 13 y las 16 palabras de longitud. Las palabras que aparecen al final de cada oración están igualadas en frecuencia de uso y longitud (en sílabas).

La prueba de vocabulario emplea el emparejamiento de palabras y dibujos y contiene los sustantivos, combinaciones sustantivo + adjetivo y los verbos en infinitivo que aparecen en la prueba de CO. Todos estos elementos son de moderada a alta frecuencia de uso [22]. El objetivo en este caso es determinar si existen problemas específicos en la conexión entre las representaciones ortográficas de las palabras y su significado.

En uno de los primeros estudios realizado con la batería, se aplicó la prueba de CO a una muestra de 29 adultos mayores [23]. La edad media de este grupo era de $72,55 \pm 8,54$ años. El nivel general de funcionamiento cognitivo se estimó con el minexamen cognitivo (MEC) [24], la escala de deterioro global (GDS) [25] y la clasificación clínica de la demencia (CDR) [26]. La capacidad de la memoria operativa verbal se evaluó con la prueba complementaria incluida en la batería.

Se realizó un análisis de *cluster* que permitió dividir la muestra en dos grupos que diferían en su capacidad de MO, y en el rendimiento ante los distractores léxicos y sintácticos de la prueba de CO. Sin embargo, los grupos estaban emparejados en variables como la edad, los años de escolaridad, la puntuación del MEC, la GDS y la CDR.

Se comprobó que el grupo de baja capacidad de MO mostraba peor ejecución que el de alta capacidad en la mayoría de estructuras oracionales. Esta diferencia se debía sobre todo al bajo rendimiento del primer grupo ante los distractores. Los resultados mostraron también que no existían efectos diferenciales de la complejidad sintáctica, pero sí del número de proposiciones, especialmente en los distractores sintácticos.

Otro ejemplo puede ilustrar el uso de la prueba de CO en la investigación sobre el envejecimiento. En un estudio más reciente [27], se investigó la

Figura 1. Ejemplos de un ítem congruente (a), de un distractor sintáctico (b) y de un distractor léxico (c).



variabilidad interindividual en la comprensión de oraciones para un total de 196 personas, de edades comprendidas entre los 50 y los 80 años, todos con estudios primarios. Se calcularon los aciertos en cuatro clases de oraciones: una proposición-orden canónico, una proposición-orden no canónico, dos proposiciones-orden canónico, dos proposiciones-orden no canónico.

El análisis de regresión de la edad sobre las medidas de la variabilidad –residuos estandarizados en valor absoluto de los índices originales– [28] mostró que la diversidad aumenta de forma lineal a medida que aumenta la edad en los tres primeros tipos de oraciones, y que en las oraciones más difíciles (dos proposiciones-orden no canónico) la variabilidad aumenta hasta los 65-66 años, pero disminuye a partir de aquí, ajustándose así a una función cuadrática con forma de U invertida. En el grupo de mayor edad, la media de aciertos en estas oraciones se aproxima al nivel de acierto por azar (10,5/18), por lo que es posible concluir que la disminución en la variabilidad a partir de los 65-66 años indica un cambio hacia un nivel de comprensión inadecuado de estas oraciones.

Este trabajo tiene dos objetivos: realizar un análisis de la prueba de CO basado en la teoría clásica de los test y en la teoría de la respuesta al ítem (TRI), y comprobar si la dimensionalidad de la prueba se mantiene en distintas etapas evolutivas.

Sujetos y métodos

Muestra

La prueba de CO se aplicó a una muestra total de 2.238 sujetos de edades comprendidas entre los 6 y

los 85 años (media: $39,18 \pm 24,71$ años), de los cuales el 44,7% eran varones, y el 55,3%, mujeres.

Con el objetivo de comprobar si la estructura dimensional de la prueba se mantiene a través de distintos períodos evolutivos, se establecieron tres grupos de edad: 6-15 años ($n = 664$; 44% varones y 56% mujeres), 16-60 años ($n = 1.034$; 44,1% varones y 55,9% mujeres) y 61-85 años ($n = 540$; 46,9% varones y 53,1% mujeres). La media de edad del primer grupo era de $10,33 \pm 2,79$; del segundo grupo, de $40,19 \pm 13,13$, y del tercero, de $72,71 \pm 6,44$ años.

Los tres tramos de edad pretenden reflejar el período de desarrollo o crecimiento de la comprensión gramatical (6 a 15 años), la etapa de madurez o estabilidad (16 a 60 años) y la fase de declive (61 a 85).

Además, esta decisión está motivada por criterios técnicos, ya que los análisis de datos pertinentes para el estudio requieren que el tamaño de cada grupo sea grande, dada la relación entre el tamaño muestral y la precisión del estimador de la fiabilidad.

Instrumentos

La prueba de CO incluye 102 elementos que pueden agruparse en 17 categorías. En cada categoría aparecen seis ejemplos: dos congruentes, dos distractores sintácticos y otros dos distractores léxicos (Fig. 1).

En la tabla I aparece cada tipo de oración, sus rasgos (el número de proposiciones que incluye y su ajuste al orden canónico del español) y un ejemplo de cada clase.

Procedimiento

En la aplicación de la prueba se presenta cada par oración-dibujo en la pantalla del ordenador hasta que el sujeto responde ‘verdadero’ o ‘falso’. Al res-

Tabla I. Tipos de oraciones, rasgos de cada tipo y ejemplos.

Tipo de oración	N.º de proposiciones	Ajuste al orden canónico	Ejemplos
Pasivas verbales	1	No	El hombre es adelantado por el caballo
Pasivas verbales-S. prep.-SN	1	Sí	Es despertado por el hombre el niño
Pasivas verbales-SN-S. prep.	1	No	Es atacado el gato por el niño
Pasivas truncadas: sujeto-verbo	1	No	El gato es perseguido
Pasivas truncadas: verbo-sujeto	1	No	Es atacado el gato
Pasivas de dativo	1	No	La flor es entregada a la mujer por el hombre
De relativo pasivizadas de sujeto	2	Sí	El perro que está arrastrando al gato es pequeño
De relativo pasivizadas de objeto	2	No	El gato que el caballo está persiguiendo es blanco
Activas	1	Sí	El caballo mordió al perro
Activas no canónicas	1	No	Al abuelo le acarició la niña
De dativo	1	Sí	El niño entregó el vaso a la abuela
De relativo anidadas de sujeto	2	Sí	El perro que mordió al caballo es grande
De relativo anidadas de objeto	2	No	El perro a quien el niño arrastró es pequeño
De relativo sujeto-objeto	2	No	El perro al que el gato mordió, empuja al niño
De relativo objeto-sujeto	2	Sí	El niño besó a la mujer que arrastra al perro
De sujeto focalizado	1	Sí	Es el perro el que mordió al gato
De objeto focalizado	1	No	Es a la mujer a la que despierta el hombre

SN: sintagma nominal; S. prep.: sintagma preposicional.

ponder con el ratón, aparece un nuevo ítem en la pantalla. Este procedimiento se repite hasta el final de la prueba. El conjunto de ítems de la prueba de CO está dividido en dos bloques de 51 elementos, lo que permite realizar un descanso en la mitad de la aplicación. No se proporciona *feedback* al sujeto en ningún momento.

Programas empleados para los análisis y criterios de valoración

Se empleó el programa TESTFACT 4.0 [29] para calcular el coeficiente de fiabilidad (KR-20) de la prueba con la muestra total y para los tres grupos. Este programa permite también comprobar la di-

mensionalidad de la prueba –global– y la específica de cada grupo.

Una prueba puede considerarse unidimensional si la *ratio* entre el primer autovalor y el segundo está próxima a cinco, de acuerdo con el criterio de Lord [30]. Este criterio debe cumplirse para poder realizar los análisis basados en la TRI. Además, si los pesos de los ítems en el factor único están por debajo de 0,3, se consideran no ajustados.

Para los análisis basados en la TRI se seleccionó el modelo de tres parámetros (discriminación, facilidad y probabilidad de acierto por adivinación). Los análisis se llevaron a cabo con el programa BILOG-MG 3 [31]. En relación con el parámetro de discriminación, un valor del estimador mayor que uno indicaría que

Tabla II. Coeficiente de fiabilidad (KR-20), primer y segundo autovalor en el análisis factorial y porcentaje de ítems fáciles o muy fáciles para cada grupo de edad.

Grupos	KR-20	1.º autovalor	2.º autovalor	Ítems fáciles o muy fáciles
6-15 años	0,953	34,76	7,06	85%
16-60 años	0,948	35,89	9,01	94%
61-85 años	0,951	29,06	8,82	80%

el ítem permite diferenciar entre sujetos que tienen una aptitud inferior a la posición que ocupa ese ítem y los que la tienen superior. El parámetro de dificultad manifiesta la posición del ítem en la escala de la habilidad latente, y se expresa en puntuaciones típicas con media igual a cero. Por último, el parámetro de probabilidad de acierto por adivinación expresa la probabilidad de acertar el ítem por parte de sujetos que tienen los niveles mínimos de la habilidad.

En la TRI se sustituye el concepto de fiabilidad por el de función de información, que manifiesta el grado en que el test resulta informativo en los diferentes valores de las puntuaciones [32]. El valor de la información depende del valor del parámetro de discriminación. El error típico de estimación está inversamente relacionado con la información, y también puede expresarse como función del nivel del rasgo que poseen los sujetos.

Resultados

El análisis global de la prueba de CO mostró que ésta puede considerarse unidimensional ($ratio = 37,7/8,76$). El coeficiente de fiabilidad KR-20 obtenido para el test fue igual a 0,956. En este análisis aparecieron seis ítems cuyos pesos en el factor único están por debajo de 0,3. Además, la mayoría de los ítems de la prueba son fáciles o muy fáciles (92%), y el resto presenta una dificultad media.

Se realizaron análisis separados para cada uno de los grupos de edad ya citados. La tabla II resume los resultados del análisis realizado para cada grupo.

Como puede observarse en la tabla II, en el grupo de 6-15 años la prueba puede considerarse unidimensional. En este análisis aparecieron 11 ítems con pesos inferiores a 0,3 en el factor único; además, el 15% tenía una dificultad media o alta, y el resto era fácil o muy fácil.

En el grupo de 16-60 años se confirma la unidimensionalidad de la prueba. En este caso, el por-

centaje de elementos que presentan pesos menores a 0,3 en el factor único alcanzó el 34%. La mayoría de estos últimos son ítems congruentes, que, por otro lado, serían los más fáciles. La mayoría de los ítems presentaban una dificultad baja o muy baja, y el resto (4%) una dificultad media.

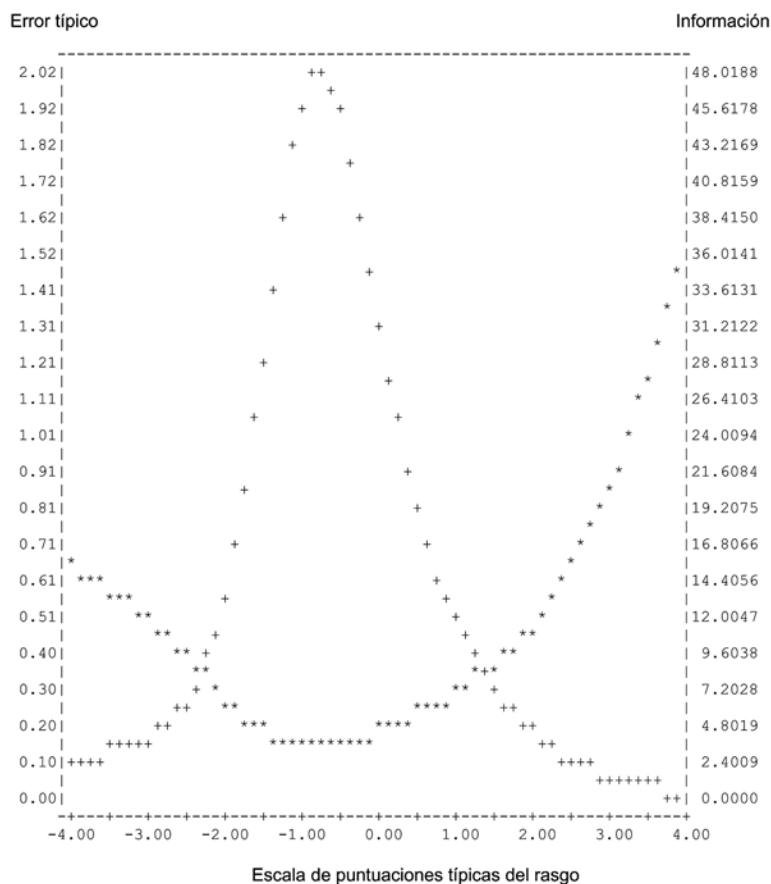
En el grupo de 61-85 años, la prueba debe considerarse bidimensional. Al analizar los elementos que formaban parte de cada uno de los factores identificados, se comprobó que el segundo factor incluía todos los ítems congruentes y el primero todos los distractores. En este caso, el 20% de los elementos presentaba una dificultad media y alta, y el resto era fácil o muy fácil.

Análisis basados en la TRI

Los resultados del análisis realizado para el grupo de 6 a 15 años indicarían que aproximadamente la mitad de los ítems muestran una discriminación adecuada, es decir, permiten diferenciar entre sujetos que tienen una aptitud inferior a la posición que ocupa el ítem y los que la tienen superior. La mayoría de estos ítems son distractores (48/52), y el resto elementos congruentes.

Los resultados obtenidos en relación con el parámetro de dificultad permiten concluir que se trata de una prueba fácil (aproximadamente, un 85% de elementos fáciles o muy fáciles). Los ítems más fáciles fueron los congruentes. Por otro lado, los resultados del análisis indicaron que 10 ítems –todos congruentes– presentaban una probabilidad de acierto por adivinación igual o aproximadamente igual a 0,5. En general, los ítems mostraron un buen ajuste al modelo de tres parámetros (excepto 11 elementos, casi todos congruentes).

En el grupo de 16 a 60 años, algo más de la mitad de los ítems (52%) mostró una discriminación adecuada. La mayoría de estos elementos eran distractores (49/53). Los ítems que presentaban los valores más bajos de discriminación eran todos congruen-

Figura 2. Funciones de información (línea de +) y de error típico (*) del test de comprensión de oraciones.

tes, y éstos resultaron también los más fáciles. La mayoría de los ítems mostró un buen ajuste al modelo de tres parámetros, aunque el porcentaje de ítems desajustados se acercaba al 20% en este grupo (casi todos vuelven a ser congruentes), los mismos ítems que en el grupo anterior presentaron probabilidades de acierto por adivinación iguales o próximas a 0,5.

Función de información y error típico de estimación

La figura 2 muestra los resultados globales obtenidos con la TRI, las funciones de información y de error típico de estimación. Se observa que la máxima información proporcionada por el test es aproximadamente de 48,02 en un valor del rasgo próximo a $-0,75$. La curva del error típico de estimación muestra que el mínimo coincide con la información

máxima (siendo aproximadamente de 0,2). El error típico aumenta especialmente cuando las puntuaciones típicas del rasgo son muy altas, a partir de 2,5 desviaciones típicas por encima de la media.

Conclusiones

La tarea de verificación simple empleada en la prueba de CO demanda pocos recursos de memoria, lo que hace que sea especialmente adecuada para su aplicación en poblaciones especiales. Por otro lado, tanto por sus características, número de estructuras sintácticas y rango de edades al que puede ser aplicada (6/7 a 85 años), es única en español.

El diseño de la prueba permite comprobar si los sujetos muestran sensibilidad al número de proposiciones –es decir, si tienen mejor rendimiento con las oraciones de una proposición que con las de dos– o a la complejidad sintáctica, como se ha demostrado en alguno de los estudios previos realizados con ella [23]. Además, pueden realizarse otros para los distintos tipos de ítems (congruentes, distractores léxicos y sintácticos). Un bajo rendimiento ante los distractores sintácticos revelaría la existencia de dificultades en la comprensión de oraciones, mientras que en el caso de los distractores léxicos señalaría una dificultad de carácter léxico-semántico. El análisis puede realizarse también por clases de oraciones [27] –combinando ortogonalmente los rasgos subyacentes de las oraciones: el número de proposiciones y el ajuste al orden canónico–.

La prueba de CO, que constituye el elemento nuclear de la batería ECCO, tiene una fiabilidad global alta, e incluye elementos fáciles o muy fáciles, lo que la hace muy sensible a las dificultades que pueda presentar un sujeto en el proceso de asignación de los roles temáticos. Por otro lado, al considerar la muestra completa, la prueba resulta unidimensional, y la mayoría de los elementos presenta pesos en el factor único que indican un buen ajuste.

En los análisis realizados para los tres grupos de edad, los coeficientes de fiabilidad se mantienen por encima de 0,94 en todos los casos. El porcentaje de ítems fáciles o muy fáciles es menor en los grupos extremos, aunque siempre iguala o supera el 80%.

La unidimensionalidad de la prueba de CO se confirma en los dos primeros grupos de edad, aunque en el segundo (16-60 años) aparece un mayor porcentaje de ítems no ajustados. Sin embargo, en el grupo de adultos mayores aparecen dos dimensiones: la primera identificada con los distractores, y la segunda con los ítems congruentes. Este resultado indicaría que el rendimiento de este grupo

ante los distractores y ante los ítems congruentes está claramente dissociado, como aparece también en alguno de los trabajos ya citados [23] cuando se utiliza como estrategia el agrupamiento de sujetos.

Los análisis basados en la TRI han puesto de manifiesto que los elementos más fáciles, menos discriminativos y peor ajustados al modelo de tres parámetros son los congruentes. Sin embargo, los ítems que mejor discriminan son los distractores.

La función de información obtenida muestra que la prueba mide la comprensión de oraciones con diferente precisión a lo largo de la escala de aptitud: proporciona información adecuada para las puntuaciones centrales, resultando poco informativa a partir de una desviación típica por encima de la media y de 2,5 desviaciones típicas por debajo de ésta. Puede afirmarse, por tanto, que la prueba está sesgada hacia niveles medios y bajos de la comprensión gramatical.

En conclusión, las propiedades psicométricas de la prueba de CO y los resultados obtenidos en algunos estudios avalan la utilidad de la batería como un instrumento de evaluación de la comprensión gramatical en niños y adultos, así como en el ámbito clínico y en la investigación. Por ello, se están desarrollando otros estudios que permitirán conocer, por ejemplo, los tipos de oraciones que mejor discriminan entre los pacientes con deterioro cognitivo y las personas mayores sin deterioro, o las posibles diferencias entre estos grupos en la codificación de los factores que permiten caracterizar las oraciones (número de proposiciones y ajuste al orden canónico).

Bibliografía

1. Valle F, Cuetos F. EPLA: evaluación del procesamiento lingüístico en la afasia. London: LEA; 1996.
2. Stark J. Everyday life activities. Wien: Druckerei Bosmuller; 1992.
3. Aguado G. El desarrollo de la morfosintaxis en el niño: TSA. Madrid: CEPE; 1989.
4. Bishop D. TROG. Test for Reception of Grammar. Medical Research Council. Manchester: Chapel Press; 1983.
5. Mendoza E, Carballo G, Muñoz J, Fresneda MD. (CEG) Test de comprensión de estructuras gramaticales. Madrid: TEA; 2005.
6. Puyuelo M, Wiig E, Renom J, Solanas A. Batería de lenguaje objetiva y criterial (BLOC). Barcelona: Masson; 1998.
7. Kemper S, Kliegl R. Limitations of syntactic processing. In Kemper S, Kliegl R, eds. Constraints on language: aging, grammar and memory. Boston: Academic Publishers; 1999. p. 79-105.
8. Wingfield A, Peelle JE, Grossman M. Speech rate and syntactic complexity as multiplicative factors in speech comprehension by young and older adults. *Aging Neuropsychol Cogn* 2003; 10: 310-22.
9. Almagro Y, Sánchez-Casas RM, García-Albea JE. El agramatismo y su sintomatología. *Rev Neurol* 2005; 40: 369-80.
10. Friedman N, Gvion A. Sentence comprehension and working memory limitation in aphasia: a dissociation between semantic-syntactic and phonological reactivation. *Brain Lang* 2003; 86: 23-39.
11. Caplan D, Waters GS. On-line syntactic processing in aphasia: studies with auditory moving window presentation. *Brain Lang* 2003; 84: 222-49.
12. Waters GS, Caplan D. Working memory and online syntactic processing in Alzheimer's disease: studies with auditory moving window presentation. *J Gerontol Ser B Psychol Sci Soc Sci* 2002; 57: 298-311.
13. Waters GS, Rochon E, Caplan D. Task demands and sentence comprehension in patients with dementia of the Alzheimer's type. *Brain Lang* 1998; 62: 361-97.
14. Bickel C, Pantel J, Eysenbach K, Schroder J. Syntactic comprehension deficits in Alzheimer's disease. *Brain Lang* 2000; 71: 432-48.
15. Hochstadt J, Nakano H, Lieberman P, Friedman J. The roles of sequencing and verbal working memory in sentence comprehension deficits in Parkinson's disease. *Brain Lang* 2006; 97: 243-57.
16. Mendoza E. Trastorno específico del lenguaje. Madrid: Pirámide; 2001.
17. Muñoz-López J, Fresneda MD, Mendoza-Lara E, Carballo G. Propiedades psicométricas de una prueba de comprensión gramatical. *Rev Neurol* 2008; 47: 21-6.
18. Crain S, Ni W, Shankweiler D. Grammatism. *Brain Lang* 2001; 77: 294-304.
19. López-Higes R, Del Río D, Fernández S. Batería ECCO (exploración cognitiva de la comprensión de oraciones). Madrid: EOS; 2005.
20. Mitchum CC, Berndt RS. Cognitive neuropsychological approaches to diagnosing and treating language disorders: production and comprehension of sentences. In Chapey R, ed. Language intervention strategies in aphasia and related neurogenic communication disorders. London: Lippincott Williams & Wilkins; 2001. p. 551-71.
21. Daneman M, Carpenter PA. Individual differences in working memory and reading. *J Verbal Learn Verbal Behav* 1981; 19: 450-66.
22. Alameda JR, Cuetos F. Diccionario de frecuencias de las unidades lingüísticas del castellano. (Vols. 1 y 2). Oviedo: Servicio de Publicaciones de la Universidad de Oviedo; 1995.
23. Martín-Aragoneses MT, López-Higes R, Fernández S, Del Río D, González J. Psycholinguistic markers of cognitive deterioration. In Ballesteros S, ed. Aging, cognition and neuroscience. Madrid: UNED; 2006.
24. Lobo A, Saz P, Marcos G. Examen cognoscitivo minimal. Madrid: TEA; 2002.
25. Reisberg B, Ferris SH, De León MJ, Crook T. The Global Deterioration Scale for assessment of primary degenerative dementia. *Am J Psychiatry* 1982; 139: 1136-9.
26. Hughes CP, Berg L, Danzinger WL. A new clinical scale for staging of dementia. *Br J Psychiatr* 1998; 140: 5666-762.
27. López-Higes R, Rubio S, Martín MT, Del Río D. Variabilidad en la comprensión gramatical en el envejecimiento normal. *Revista de Logopedia, Foniatría y Audiología* 2008; 1: 9-21.
28. Hultsch DE, MacDonald WS, Dixon RA. Variability in reaction time performance of younger and older adults. *J Gerontol Psychol Sci* 2002; 57: 101-15.
29. Wood R, Wilson D, Gibbons R, Schilling S, Muraki E, Bock D. TESTFACT 4. Scientific Software International. Michigan: Lincolnwood; 2003.
30. Lord FM. Applications of item response theory to practical testing problems. Hillsdale, NJ: Lawrence Erlbaum; 1980.
31. Zimowsky M, Muraki E, Mislavsky R, Bock D. BILOG-MG 3. Scientific Software International. Michigan: Lincolnwood; 2003.
32. Martínez-Arias R. Psicometría: teoría de los test psicológicos y educativos. Madrid: Síntesis; 1995.

Reliability and applications of a grammar comprehension assessment test

Introduction. The sentence comprehension test from the sentence comprehension cognitive examination (ECCO) battery can be used to assess children and adults with grammar comprehension disorders by means of a simple verification procedure that requires few memory resources.

Aim. To present a psychometric analysis of the sentence comprehension test based on the classical theory of tests and on item response theory, as well as several studies that show its usefulness for assessing grammar comprehension in different populations.

Subjects and methods. The test was applied to 2238 subjects with ages ranging from 6 to 85 years.

Results. The analysis shows that the test has a high overall reliability and, through the age groups that were considered, it includes elements that are especially discriminating or informative (lexical and syntactic distractors) and is biased towards medium-low levels of the trait.

Conclusions. In view of the results, the sentence comprehension test from the ECCO battery provides a reliable measure of grammar comprehension through a wide variety of constructions and has great potential value for use in the clinical and research domains.

Key words. Comprehension disorders. Grammar comprehension. Item response theory. Lexical distractor. Reliability. Syntactic distractor. Written language assessment.