

EL TEST CLOZE EN LA EVALUACION DE LA COMPRESION DEL TEXTO INFORMATIVO DE NIVEL UNIVERSITARIO

CLOZE TEST IN ASSESMENT OF READING COMPREHENSION OF EXPOSITORY TEXT AT COLLEGE LEVEL

HILDA DIFABIO DE ANGLAT
Centro de Investigaciones Cuyo – CONICET
ganglat@gmail.com

RESUMEN

En una muestra de 114 estudiantes de Ciencias de la Educación, Universidad Nacional de Cuyo, con un texto preparado *ad hoc* se evalúa la potencialidad del *test de cloze* para evaluar la comprensión de un texto informativo y se relacionan sus resultados con el rendimiento académico. El índice medio de dificultad de los ítems del *cloze* es 0,67; luego, resulta ligeramente más sencillo que lo recomendado por la teoría estadística para el ámbito educativo (entre 0,5 y 0,6). Alcanza una correlación significativa de 0,51 con el rendimiento académico.

Palabras claves: Comprensión de texto, dificultad, nivel universitario, rendimiento académico, test de cloze.

ABSTRACT

In a sample of 114 students of Educational Sciences, Universidad Nacional de Cuyo, cloze test's potentiality to assess reading comprehension of expository text is measured and correlation with academic performance is established. The median difficulty index is 0.67, so it results slightly easier than statistical theory recommendation (between 0.5 and 0.6) for educational settings. It reaches a meaningful correlation of 0.51 with academic performance.

Keywords: Reading comprehension, difficulty, college level, academic performance, cloze test.

Recibido: 10-11-2007. *Aceptado:* 25-03-2008.

1. INTRODUCCION

EN LA comprensión del texto se halla implicada toda una gama de operaciones y habilidades intelectuales que confluyen para producir un logro intelectual

fundamental: la captación de una organización de significados de un texto escrito, formato bajo el cual éstos se presentan en el nivel universitario. Dada la importancia de la lectura, sobre todo en este nivel de la educación formal, es necesario implementar instancias ponderativas que permitan a los alumnos conocer sus potencialidades así como las limitaciones que deben superar en pro de buen desempeño académico. Para ello, dichas instancias deberán respetar una serie de requisitos: emplear textos auténticos –en los que el conocimiento previo es una variable importante del rendimiento–, que promuevan la *inferencia* –componente esencial de la comprensión, aun a nivel oracional–, la síntesis, la habilidad para adaptar las estrategias lectoras al texto o a la situación, el empleo flexible del conocimiento en situaciones nuevas, entre otros.

En la literatura especializada se advierte hoy, en distintas lenguas, un resurgimiento del interés por el *test de cloze* (cfr., por ejemplo, Cheng, 2004; Dastjerdi y Talebinezhad, 2006; Greene, 2001; Kobayashi, 2002, 2004; Oliveira y Santos, 2006; Yan, 2006), instrumento diseñado por Wilson Taylor (1953) para evaluar la legibilidad¹ del texto. En este sentido, se utiliza en un número significativo de trabajos para determinar niveles de logro en comprensión de textos expositivos en el nivel universitario, punto de partida imprescindible en un diagnóstico, puesto que, como señaláramos, es principalmente mediante lecturas como los estudiantes acceden al contenido de las diversas disciplinas.

En este trabajo presentamos, en primer lugar, cuatro marcos conceptuales complementarios que explican los procesos explorados por el *cloze*. En segundo lugar, los resultados de la aplicación de un texto *cloze* preparado *ad hoc* sobre un tema de metodología de la investigación (el diseño) a estudiantes de tercer año de Ciencias de la Educación en cursos sucesivos (desde el 2003 al 2006), centrándonos en el nivel de dificultad de los ítemes y los patrones de errores; finalmente, se correlacionan los resultados en el *cloze* con el rendimiento académico.

2. ENCUADRE TEORICO

El procedimiento *cloze*² es la mejor medida que conocemos para evaluar los procesos esenciales de la lectura (anticipación, inferencia, juicio, resolución de problemas) y, por ello, resulta particularmente efectivo para detectar las dificultades del

¹ Es una de las traducciones del término inglés *readability*; Rodríguez Diéguez (1991) y Morles (1981), divulgadores del procedimiento en América Latina, lo traducen como lecturabilidad y comprensibilidad, respectivamente.

² El término *cloze* parece ser un neologismo, pues no figura en los diccionarios (Cairney, 1996: 95, nota del traductor).

alumno, promover su advertencia metacognitiva de las mismas y emprender acciones de reparación.

El *cloze* estándar consiste en un texto al que se le ha suprimido una palabra cada cinco, dejando intactas la primera y la última oración (o, en otros autores –González Moreyra, 1998–, las diez primeras y las diez últimas palabras del texto), a fin de que el lector “llene” los blancos.

Es una técnica eficaz, tanto desde el punto de vista práctico (es fácil de confeccionar, aplicar y calificar) como empírico en razón del alto índice de correlación positiva que evidencia con la evaluación del desempeño académico (coeficientes entre 0,46 y 0,72 según el área de estudio y la edad, cfr. Oliveira y Santos, 2006: 22). También en nuestro estudio de regresión entre variables (nivel intelectual, pensamiento crítico y comprensión lectora) que explican el rendimiento en nivel medio y universitario, la comprensión lectora –evaluada mediante un texto *cloze*– explica casi el 40% de la varianza (Difabio de Anglat, 2005). Es una asociación que no sorprende ya que hemos “medido” el rendimiento con las calificaciones y éstas con frecuencia reflejan la retención y, a veces, cierto grado de elaboración de la información vehiculizada en los textos.

Desde el punto de vista teórico, cuatro marcos conceptuales complementarios explican los procesos explorados por el *cloze*.

En la propuesta originaria de Taylor, el *test de cloze* se define por el principio de cierre gestáltico (*closure*) de la Psicología de la forma; el autor acuñó el término a partir de este principio (Artola González, 1991: 69): la tendencia de los sujetos a “completar un patrón familiar no lo bastante acabado, a ver un círculo roto como uno entero (...) cerrando mentalmente el hueco” (Taylor, 1953: 415). Esto es, ya que una buena forma es una *gestalt* o forma cerrada, el lector tiende a “cerrar” las percepciones incompletas mediante el conocimiento previo, la inferencia y la expectativa intuitiva³: “Dada la secuencia ‘el viejo... a lo largo de la polvorienta carretera’, casi todo el mundo la completaría con alguna forma verbal (disposición gramatical), y la mayoría se vería afectada por el elemento semántico ‘viejo’ y elegiría un verbo apropiado tal como ‘rengueaba’, ‘se arrastraba’, ‘cojeaba’” (Osgood y Sebeok, 1965: 158).

Desde la teoría sistémica, se lo relaciona con el concepto de *entropía* o incertidumbre creada por la falta de organización de un sistema, que se opone al concep-

³ “Las figuras incompletas tienden a percibirse como totalidades cerradas. Por eso, al leer entendemos el significado de una palabra estando ausente una sílaba, o se comprende una oración aunque falte una palabra. Véanse estos dos enunciados ilustrativos de cada alternativa:

Miguel de Cerantes escribió la novela Don Quijote de la Mancha.

Miguel de escribió la novela Don Quijote de la Mancha.

En el primero, un individuo de habla española leerá CERVANTES y no CERANTES; en el segundo, leerá MIGUEL DE CERVANTES (...) y no MIGUEL DE ESCRIBIÓ (...)” (Navarro, 1991: 68).

to de redundancia⁴; luego, se propone su definición en términos de lectura de dicha redundancia: el buen lector es capaz de leer la información distribuida en el texto y por ello puede inferir los términos elididos. Para ilustrar con un ejemplo sencillo, el género de una unidad léxica, en una construcción nominal corriente, es conducido usualmente por –al menos– tres vocablos: el artículo, el sustantivo y el adjetivo.

En tercer lugar, se lo vincula con el análisis del discurso ya que requiere que el lector movilice todos sus recursos lingüísticos. Por lo tanto, no es un simple ejercicio de completamiento que manifieste habilidades lectoras de bajo nivel, sino una clase de análisis discursivo que implica el uso de la lengua como un todo integrado, “(...) que pone en juego la potencialidad creativa del lenguaje del sujeto en un proceso de composición de carácter convergente” (González Moreyra, 1998: 48).

Finalmente, Condemarín y Milicic (1988) señalan que guarda relación estrecha con el enfoque psicolingüístico que caracteriza la lectura como un proceso interactivo entre el procesamiento de los datos aportados por el texto (proceso de abajo hacia arriba) y la anticipación que realiza el lector (proceso de arriba hacia abajo), quien incrementa su habilidad para obtener información de la lectura gracias al éxito de sus predicciones. Para completar la palabra omitida, forma hipótesis y predicciones mediante procesos inferenciales a partir de su conocimiento previo del tópico, de su conocimiento lingüístico (de las claves fonográficas, semánticas, sintácticas, y de las restricciones del lenguaje) y de la redundancia textual. En el proceso de abajo hacia arriba, atiende a las claves disponibles en el texto (presentes en la memoria de trabajo) para confirmar, rechazar o modificar sus hipótesis; esto es, el desempeño en este proceso interactivo moviliza la conciencia lingüística o metalingüística (sensibilidad hacia la estructura interna del lenguaje) y exige la regulación metacognitiva de la tarea. Por otra parte, la motivación y el interés apoyan los subprocesos mencionados porque sostienen la atención necesaria para el procesamiento de las claves textuales.

Para interpretar los resultados, la práctica corriente es convertir los puntajes en *niveles funcionales de lectura* (Condemarín y Milicic, 1988): independiente, instruccional y de frustración.

Nivel independiente significa que el alumno lee con fluidez, precisión y comprende la mayor parte del texto.

En el *nivel instruccional*, la lectura es medianamente fluida porque aparecen algunas dificultades en el reconocimiento de palabras y, aunque el lector capta el

⁴ Proporciona un valor para cada blanco que varía desde 0% –todos los sujetos dan la misma respuesta– a 100% –cada uno propone una respuesta diferente–. El primer extremo significa máxima organización del sistema y una probabilidad de 1 (certeza) de elección de la alternativa; con un puntaje del 100% cuando la respuesta es correcta y del 0% cuando es incorrecta. El segundo manifiesta inexactitud máxima –una alternativa para cada sujeto–, lo que sugiere una alta probabilidad de predicción al azar.

contenido y su estructura, evidencia algunas fallas en la comprensión. Luego, si bien el material no le resulta sencillo, aún puede manejarlo.

El *nivel de frustración* implica que son numerosos los errores de reconocimiento de palabras y que la comprensión es ciertamente deficiente.

Los índices establecidos para el español (Condemarín y Milicic, 1988) son: 75% o más ubican al lector en el nivel independiente; entre 74 y 44%, en el instruccional; 43% o menos, en el de frustración.

González Moreyra subcategoriza los niveles de la siguiente forma (1998: 52): en el independiente, distingue el rendimiento *excelente* (100-90%) y el *bueno* (89-75%); en el nivel intermedio, que llama “dependiente”, el *instruccional* (74-58%) y el *de dificultad* (57-44%); en el “deficitario”, el *malo* (43-30%) y el *pésimo* (29-0%).

El procedimiento más frecuente en la calificación del *cloze* es aceptar sólo la palabra correcta (*verbatim*) porque la evidencia empírica muestra que el cómputo de sinónimos no altera significativamente el nivel de comprensión lectora del sujeto y es menos confiable en tanto depende del criterio subjetivo del evaluador. Sin embargo, dicho cómputo arroja correlaciones levemente más altas con otras pruebas de comprensión lectora, con alguna incidencia, por lo tanto, sobre la validez. Además, si se emplea el llamado método exacto de puntuación (en el que únicamente se acepta el término original elidido), el *cloze* resulta con frecuencia demasiado difícil incluso para hablantes adultos educados. Por otra parte, dado que se busca evaluar la capacidad del lector para captar el significado del texto, aunque no logre anticipar exactamente su estructura de superficie, se justifica un criterio más “blando” (y luego el tiempo *adicional* que ciertamente implica): clasificar las respuestas en correctas –esperadas y otras alternativas– e incorrectas.

3. ESTUDIO DE CAMPO

3.1. Participantes

Participaron 114 estudiantes que cursaron la cátedra “Métodos y técnicas de investigación educativa”, de 3^{er} año del Profesorado en Ciencias de la Educación, Facultad de Filosofía y Letras, Universidad Nacional de Cuyo, entre 2003 y 2006. La mayoría (el 93%) es de sexo femenino.

3.2. Material

Utilizamos un texto sobre diseño de investigación, elaborado *ad hoc*, estructurado con una combinación de dos criterios para la determinación de sus 60 lagunas:

razón fija (omisión del quinto vocablo) y racional. Sólo la primera oración se mantuvo íntegra (el texto se incluye en Apéndice N°1).

Se atribuye un punto a cada laguna correctamente resuelta (*verbatim*—cfr. Apéndice N° 2— o un sinónimo); luego, el total posible es 60 puntos, puntuación que se transforma en porcentaje.

3.3. Procedimiento

El instrumento se aplicó en una sesión colectiva, en situación regular de clase como parte del diagnóstico de la asignatura correlativa (“Fundamentos de investigación científica”); luego, se elige un tema ya abordado, aunque sintéticamente, por la influencia del conocimiento previo en la comprensión lectora. En este sentido, tanto en investigaciones tempranas (por ejemplo, Cohen, 1975) como recientes (Santos *et al.*, 2002), se encontraron diferencias en el desempeño en el *cloze* relacionadas con el tópico textual, porque la familiaridad con el asunto obviamente aumenta la probabilidad de acierto, “introduciendo una variable interviniente en la comparación de los desempeños” (Santos *et al.*, 2002: 550).

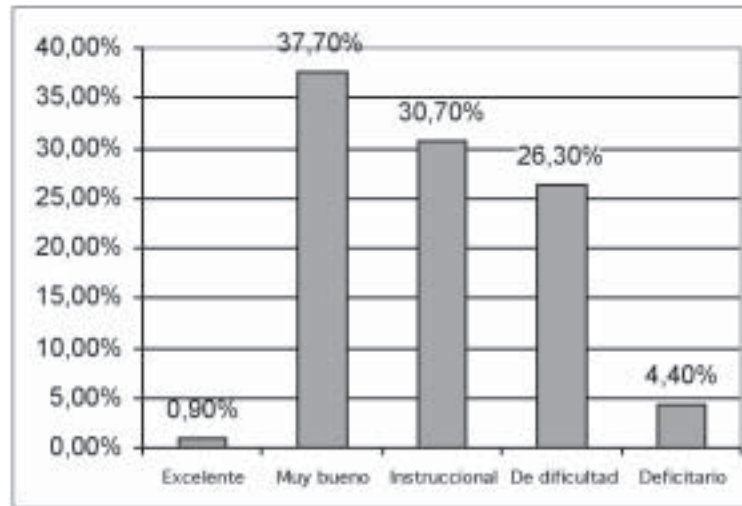
3.4. Resultados

En primer lugar, se presenta la descripción de los resultados y, luego, el análisis del nivel de dificultad de los ítems y la correlación de la puntuación total del *cloze* con el rendimiento académico.

3.4.1. Descripción de los resultados

De modo semejante a González Moreyra (1998) empleamos los siguientes criterios para la conversión de los puntajes *cloze* en niveles funcionales de lectura: 1º) Independiente, puntajes de 75% o más, subcategorizado en *excelente* (100-90%) y *muy bueno* (89-75%). 2º) Instruccional, entre 60 y 72% (porque el puntaje 74% no aparece en esta distribución). 3º) De frustración, 58% o menos (no se manifiesta el porcentaje 59), subcategorizado en *de dificultad* (58-44%) y *deficiente* (43-30%). En su representación gráfica:

Gráfico 1. Porcentajes por niveles funcionales de lectura.



Como manifiesta el gráfico de barras, un porcentaje importante de la distribución (el 69,30%) puede, al menos, “manejar” el texto.

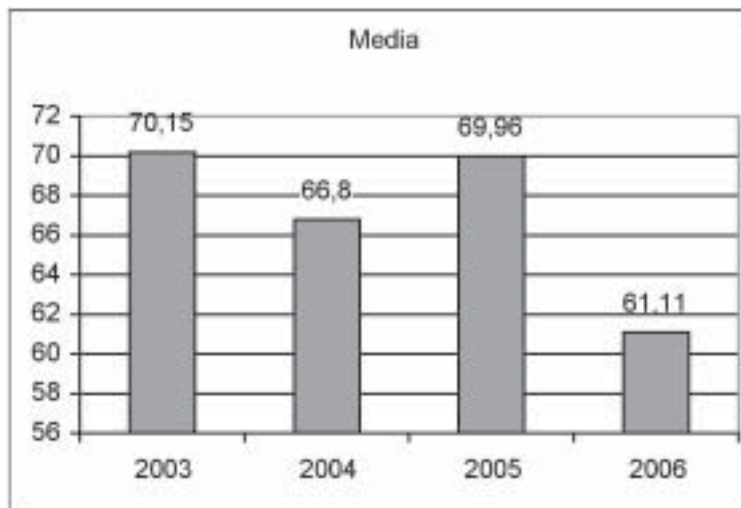
Por año, los resultados son:

Tabla I. Puntuaciones en el *test de cloze* (en porcentaje) por curso lectivo.

Curso lectivo	2003	2004	2005	2006
Casos	33	34	20	27
Media	70,15	66,80	69,96	61,11
Desviación típica	11,83	15,65	10,07	13,28
Mediana	75	71,80	71,60	60
Puntaje mínimo	47	31,60	45	37
Puntaje máximo	87	91,60	88,30	85

En su representación gráfica:

Grafico 2. Distribución del valor de la media en el *test de cloze* por curso lectivo.



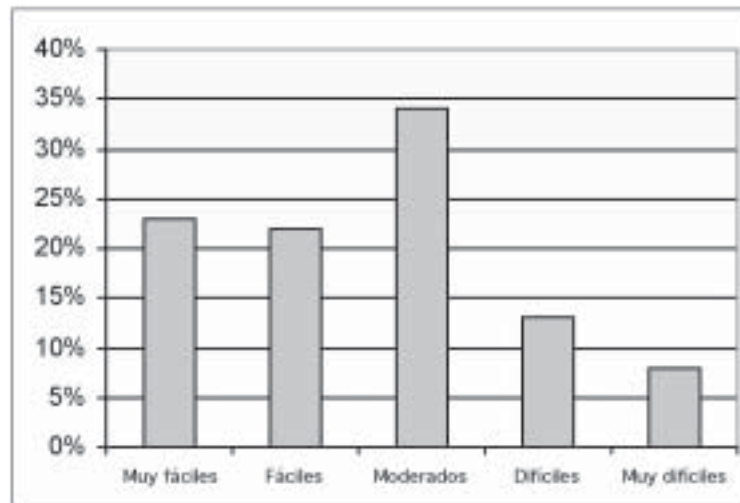
Resulta evidente un ligero decrecimiento desde el ciclo lectivo 2003 al 2004 (media de 70,15 vs. 66,8) y mayor respecto del 2006 (media de 61,11).

3.4.2. Análisis del nivel de dificultad de los ítems

Se efectúa el análisis del nivel de dificultad (*ND*) de cada ítem –una proporción de aciertos en el ítem– que se calcula por un procedimiento simple de dividir el número de alumnos que contestó correctamente el ítem entre el número total de la muestra.

El índice medio de dificultad es 0,67; luego, resulta ligeramente más sencillo que lo recomendado por la teoría estadística para el ámbito educativo porque dicho nivel medio debe oscilar entre 0,5 y 0,6. Respecto de la distribución del intervalo de dificultad de los ítems, dicha teoría aconseja un 10% de ítems muy fáciles (*ND* de 0,87 a 1), 20% fáciles (*ND* de 0,74 a 0,86), 40% con una dificultad media o moderados (*ND* de 0,53 a 0,73), 20% difíciles (*ND* de 0,33 a 0,52) y 10% muy difíciles (*ND* de 0 a 0,32). En nuestro caso, la representación gráfica de los cinco niveles de dificultad de los ítems muestra la siguiente distribución:

Gráfico 3. Porcentaje de los niveles de dificultad de los ítems.



El gráfico muestra que la distribución de los porcentajes resulta sesgada hacia los ítems de mayor facilidad: los ítems muy fáciles superan el porcentaje recomendado (representan el 23% del *cloze*), los fáciles alcanzan el criterio (22%), los moderados se acercan (34%), los difíciles están por debajo (13%) y los muy difíciles, bastante cercanos (8%). En suma, los ítems se agrupan hacia la izquierda del *continuum* de dificultad: el 45% del *test* es fácil-muy fácil *vs.* 21% difícil-muy difícil.

A partir del análisis de Abraham y Chapelle (1992), desestimando los factores extrínsecos⁵ que inciden en la dificultad de un ítem, entre los factores intrínsecos⁶

⁵ Incluyen el conocimiento del sujeto sobre el contenido del ítem y las estrategias que usa para completarlo, factores que exigirían la aplicación del *cloze* en forma individual con la técnica de pensar en voz alta.

⁶ Definen los factores intrínsecos como los elementos de contenido de la tarea “que pueden ser especificados por medios objetivos u operacionales sin referencia a las características del sujeto” (Abraham y Chapelle, 1992: 469): el contenido del ítem, los procesos cognitivos básicos requeridos para completar el ítem y el tipo de respuesta necesaria. En el primero, incluyen: la cantidad de cotexto requerido para completar el ítem (cotexto inmediato, intra-cláusula, inter-cláusula e inter-oraciones es la progresión de la dificultad) y el número de ocurrencias de la palabra correcta en el texto (análisis que interesa sólo para las palabras de contenido). El segundo factor –los procesos cognitivos requeridos para restaurar el término sobre la base de pistas semánticas y sintácticas– se operacionaliza en: palabra de contenido o función, la longitud de la palabra que debe recuperarse, el número de respuestas posibles y el número de formas posibles (las palabras que requieren un morfema de inflexión son más difíciles que las palabras que tienen formas invariantes). Desestimamos el factor “tipo de respuesta necesaria” (que en los autores es “construida” *vs.* “elegida a partir de alternativas”) por dos razones: 1) aplicamos un *cloze* tradicional y 2) se aceptan los sinónimos.

nueve de los once blancos difíciles-muy difíciles de las palabras de contenido⁷ se explican por la ausencia de ocurrencias en el texto del término correcto⁸, si bien en todos los casos hay más de una respuesta posible. En los restantes dos⁹, aunque el término aparece una vez en el texto, su localización es muy alejada del ítem.

El análisis de las discrepancias aporta criterios para interpretar cualitativamente los resultados ya que el error raramente es una respuesta al azar: refleja insuficientes destrezas lingüísticas (sintácticas y morfológicas) o fallas en el análisis lógico y proceso inferencial que llevan a desatender las claves semánticas (entre ellas, la redundancia textual). Esto es, permite distinguir entre las buenas discrepancias, aquellas que preservan el sentido –por ej., error en la concordancia de género– y discrepancias que alteran el significado.

Este análisis ha puesto de manifiesto los siguientes patrones de errores: inadvertencia de la redundancia textual¹⁰ y su contrapartida, pseudo-redundancia textual¹¹; fallas en el análisis lógico¹²; errores en el proceso inferencial a partir de un empleo abusivo del conocimiento previo o esquema¹³; uso de proformas, sobre todo nominales, que –en una analogía con la sustitución– pueden considerarse *hipernóminos* característicos del texto informativo (que “valen” para casi cualquier texto, pero lesionan la cohesión y no conducen al significado)¹⁴.

Aunquese han señalado limitaciones al *cloze* como medida de la comprensión lectora (entre otros, Gagatsis, 1985)¹⁵, la investigación ha mostrado que es una

⁷ La predictibilidad de una palabra es inversamente proporcional al tamaño de su clase: las palabras léxicas o de contenido, dado que pertenecen a clases abiertas, son más difíciles de predecir y, como contienen mayor información, un *cloze* que ha omitido una alta proporción de ellas tiene mayor entropía y luego es menos legible que uno en su versión original y, aún más, que uno que suprime principalmente palabras estructurales.

⁸ Ítems N° 1 –*estrategia*–, N° 3 –*prueba*–, N° 5 –*esencia*–, N° 20 –*número*–, N° 22 –*recolectan*–, N° 29 –*observación*–, N° 30 –*trata*–, N° 40 –*usualmente*– y N° 53 –*acorde*–.

⁹ Ítems N° 38 –*longitudinal*– y N° 56 –*deliberadamente*–.

¹⁰ Por ejemplo, el blanco N° 14 –*fenómeno*– evidencia errores (ej.: *sujeto* es una resolución que aparece en las cuatro muestras por año o *comportamiento*, en tres de las muestras) a pesar de que hay varias otras respuestas posibles (ej.: *objeto*, *hecho*, *dato*). O el blanco N° 21 –*tiempo*– cuando está precedido por “puntos en el...” y en el cotexto posterior aparece el blanco *punto* seguido de “en el tiempo”.

¹¹ Por ejemplo, la elección del término “hipótesis” (que aparece tres veces en el cotexto) para completar el ítem N° 2 –*pregunta*–.

¹² Por ejemplo, la elección de los términos *efectos* o *consecuencias* para el blanco N° 31 –*factores*–.

¹³ Un resultado que sorprende son los errores en la resolución del blanco N° 13 –*variables*– por el empleo erróneo de otros términos técnicos (*hipótesis*, *muestras*) ya que, si bien no tiene formas alternativas, aparece cuatro veces en el cotexto.

¹⁴ En el ítem N° 9 –*diversas*– (con formas alternativas: *distintas*, *varias*) aparecen los errores *algunas*, *ciertas*.

¹⁵ Su mayor sensibilidad al nivel educativo que a la dificultad del texto; medición de la comprensión de grupos de palabras más que del texto como unidad; evaluación de la capacidad para reconocer modelos familiares de expresión más que ponderación de la competencia lectora; contaminación entre comprensión y producción lingüística.

técnica eficaz para determinar un índice global de comprensión y su eficacia reside en su carácter de medida procesual de la comprensión (Artola González, 1988): cada blanco es una *ventana* al proceso lector¹⁶ que, por otra parte, evita el problema de producir un lenguaje externo al texto en preguntas y respuestas anexas.

McKenna y Layton (1990) buscan aclarar la evidencia empírica conflictiva respecto de la sensibilidad del *cloze* a la habilidad del lector para integrar información entre oraciones, ya que en algunos trabajos no se advierte una diferencia significativa en los puntajes medios del *cloze* entre la versión original y una desordenada del mismo texto (metodología de Shanahan, Kamil y Webb Tobin, 1982; estudio seminal en la investigación de la validez de la técnica como medida de comprensión inter-oracional¹⁷). Se halla que, cuando se controla la habilidad lectora, el rendimiento en un *cloze* tradicional es significativamente superior en la versión original¹⁸; por ende, concluyen que la ejecución del *cloze* correlaciona con la comprensión interoracional, aunque ésta cumpla un papel mínimo (en tanto la resolución de los blancos descansa principalmente en información intraoracional). Esto es, si bien no es una medida *directa* de comprensión interoracional, opera como un buen sustituto (Gillingham y Garner, 1992). En este sentido, Kletzien (1991: 82) muestra que alumnos del ciclo superior de la escuela media “(...) en el nivel independiente, emplean un contexto amplio para construir la comprensión; en el instruccional, se centran preferentemente en el párrafo; en el de frustración, se concentran en una base más limitada, de dos o tres palabras”. Por su parte, Greene (2001: 95) concluye: “La investigación previa ha establecido la validez del *test de cloze* para evaluar la comprensión de variados tipos de textos, pero ha hecho poco respecto de examinar su adecuación para el discurso preciso y altamente conectado de las disciplinas teóricas del nivel universitario. Aunque el formato *cloze* naturalmente evalúa habilidades de inferencia, no estima automáticamente la comprensión de la macroestructura del texto. El grado en el que el *cloze* evalúa coherencia global depende de la estrategia de eliminación empleada”.

¹⁶ Una de las críticas más frecuentes a la investigación cognitiva es el empleo de evaluaciones de resultado, esto es, obtenidas después de la tarea (y por ende dependientes de la memoria) para inferir procesos que se verifican durante la misma (Baker y Brown, 1984: 23). La tarea define el propósito y, por extensión, las operaciones para lograrla; estructura la actividad, aun cuando se especifique después. Las investigaciones de proceso (que informan sobre lo que el sujeto hace **durante** la tarea) permiten explorar las operaciones subyacentes.

¹⁷ En la versión secuenciada del texto sobre un tópico desconocido, estudiantes universitarios obtienen una media ligeramente superior (y sin significación estadística) en comparación con el texto desordenado. Este resultado no prueba, sin embargo, que el *cloze* no evalúe integración interoracional porque una diferencia pequeña puede cambiar el nivel funcional de lectura de un sujeto.

¹⁸ Aunque los autores no analizan el factor *tipo de texto* y no podemos determinar si la diferencia es estadísticamente significativa, se advierte la mayor sensibilidad del *cloze* narrativo a la comprensión interoracional según lo muestra la media de los puntajes: 4,6 para la versión desordenada y 7,6 para la versión normal (una diferencia de 3 puntos) *vs.* 3,5 y 5,4, respectivamente, en el texto informativo (una diferencia de 1,9).

A nuestro juicio, las deficiencias del lector en la anticipación de la *macroestructura* explican los errores inferenciales, porque el término que emplea el sujeto puede ser “válido” desde la relación intra-oracional, pero inaceptable desde aquélla¹⁹.

3.4.3. Correlación rendimiento académico y *test cloze*

El desempeño académico se evalúa mediante el promedio de calificaciones en los exámenes finales rendidos al momento de completar el *cloze*. El índice resultante es de 0,51, significativo a nivel $p > 0,01$.

Puede objetarse que las calificaciones no expresan el real desempeño del alumno, pero dado que ninguna evaluación puede abarcar todas las posibilidades de ponderación del rendimiento, consideramos a las calificaciones como una “expresión tan parcial como otras” (Oliveira y Santos, 2006: 20), útil en la medida en que refleja con cierta objetividad el rendimiento habitual de un estudiante en una evaluación, más allá de las excepciones por interés en alguna/s materia/s específica/s, características de la relación docente-discente, circunstancias particulares de un examen puntual.

Se logra, entonces, un resultado similar al que obtienen Oliveira y Santos (2005), en estudiantes de las carreras de Administración ($r=0,54$, $p < 0,001$) y Derecho ($r=0,52$, $p < 0,001$), pero más alto al correspondiente a Psicología ($r=0,45$, $p < 0,001$).

4. CONSIDERACIONES FINALES

El *test de cloze* se ha mostrado como un instrumento adecuado para evaluar la comprensión de un texto informativo en el nivel universitario, confirmando la investigación previa sobre la eficacia de la técnica (entre otros: Abraham & Chapelle, 1992; Dastjerdi y Talebinezhad, 2006; Kletzien, 1991). Como señala Kletzien (1991: 72), los procesos intelectuales implicados en la técnica (anticipación, razonamiento, evaluación, juicio, resolución de problemas) son los mismos que son esenciales para la lectura. En nuestro caso, aparece como potencialmente útil para identificar alumnos “en riesgo” de fracaso y, de este modo, efectuar las derivaciones pertinentes a los servicios de orientación educacional con que cuenta nuestra universidad.

Ciertamente se requieren estudios cualitativos para explorar mejor la relación entre comprensión de la lectura, evaluada mediante textos *cloze*, y rendimiento académico.

¹⁹ Por ejemplo, el término modal incorrecto en el blanco N° 5 es: (...) lo que constituye la *base* de la investigación empírica (...).

REFERENCIAS

- Abraham, Roberta y Chapelle, Carol. 1992. "The Meaning of Cloze Test Scores: An Item Difficulty Perspective", en *The Modern Language Journal* 76, 468-479.
- Artola González, Teresa. 1988. "El procedimiento cloze como medida de la comprensión lectora", en *Revista Española de Pedagogía* 46(180), 351-354.
- Artola González, Teresa. 1991. "El procedimiento cloze: una revisión general", en *Revista Complutense de Educación* 2(1), 69-81.
- Baker, Linda y Brown, Ann. 1984. "Cognitive monitoring in reading", en J. Flood (Ed.), *Understanding reading comprehension: Cognition, language and the structure of prose*. Newark: IRA, 21-44.
- Cairney, Trevor. 1996. *Enseñanza de la comprensión lectora*. Madrid: Morata.
- Cohen, J. 1975. "The effect of content area material on Cloze test performance", en *Journal of reading* 19(3), 247-250.
- Condemarín, Mabel y Milicic, Neva. 1988. *Test de cloze. Aplicaciones psicopedagógicas*. Santiago de Chile: Andrés Bello.
- Cheng, Liang. 2004. "On text structure, language proficiency, and reading comprehension test format interactions: A replay to Kobayashi, 2002", en *Language Testing* 21(2), 228-234.
- Dastjerdi, Hossein y Talebinezhad, Mohammad. 2006. "Chain-preserving deletion procedure in cloze: a discoursal perspective", en *Language Testing* 23(1), 58-72.
- Difabio de Anglat, Hilda. 2005. Competencias para la comprensión de textos y el pensamiento crítico en el nivel medio y universitario. Tesis Doctoral, Mendoza, Argentina: Universidad Nacional de Cuyo.
- Gagatsis, Athanassios. 1985. "Questions soulevés par le test de closure", en *Revue Française de Pédagogie* 70, 41-50.
- Gillingham, Mark y Garner, Ruth. 1992. "Readers' Comprehension of Mazes Embedded in Expository Texts", en *The Journal of Educational Research* 85(4), 234-241.
- González Moreyra, Raúl. 1998. "Comprensión lectora en estudiantes universitarios iniciales", en *Persona* 1, 43-65.
- Greene, Benjamín. 2001. "Testing reading comprehension of theoretical discourse with cloze", en *Journal of Research in Reading* 24(1), 82-98.
- Kletzien, Sharon. 1991. "Strategy use by good and poor comprehenders reading expository text of differing levels", en *Reading Research Quarterly* 26(1), 67-84.
- Kobayashi, Miyoko. 2002. "Cloze Test Revisited: Exploring Item Characteristics with Special Attention to Scoring Methods", en *The Modern Language Journal* 86, 571-586.
- Kobayashi, Miyoko. 2004. "Investigation of test methods effects: text organization and response format: a response to Chen, 2004", en *Language Testing* 21(2), 235-244.
- McKenna, Michael y Layton, Kent. 1990. "Concurrent validity of cloze as a measure of intersentential comprehension", en *Journal of Educational Psychology* 82(2), 372-377.

- Morles, Armando. 1981. "Medición de la comprensibilidad de materiales escritos mediante pruebas *cloze*", en *Lectura y Vida. Revista Latinoamericana de Lectura* 2(4), 16-18.
- Navarro, Armando. 1991. "Procesos perceptivos y atencionales durante la lectura", en A. Puente (Dir.). *Comprensión de la lectura y acción docente*. Madrid: Pirámide/Fundación Germán Sánchez Ruipérez, 41-72.
- Oliveira, Katya y Santos, Acácia. 2005. "Compreensão em Leitura e Avaliação da Aprendizagem em Universitários", en *Psicologia: Reflexão e Crítica* 18(1), 118-124.
- Oliveira, Katya y Santos, Acácia. 2006. "Compreensão de textos e desempenho académico", en *PSIC. Revista de Psicologia da Vêtor Editora* 7(1), 19-27.
- Osgood, Charles y Sebeok, Thomas. 1965. *Psicolingüística. Problemas teóricos y de investigación*. Barcelona: Planeta.
- Rodríguez Diéguez, José Luis. 1991. "Evaluación de la comprensión de la lectura", en A. Puente (Dir.). *Comprensión de la lectura y acción docente*. Madrid: Pirámide/Fundación Germán Sánchez Ruipérez, 301-345.
- Santos, Acácia, Primi, Ricardo, Taxa, Fernanda y Vendramini, Claudette. 2002. "O teste de Cloze na avaliação da compreensão em leitura", en *Psicologia: Reflexão e Crítica* 15(3), 549-560.
- Shanahan, Timothy, Kamil, Michael y Webb Tobin, Aileen. 1982. "Cloze as a measure of intersentential comprehension", en *Reading Research Quarterly* 17(2), 229-255.
- Taylor, Wilson. 1953. "Cloze procedure: a new tool for measuring readability", en *Journalism Quarterly* 415-433.
- Yan, Ruixia. 2006. "On 'Chain-preserving deletion procedure in cloze': a replay to Dastjerdi, and Talebinezhad, 2006", en *Language Testing* 23(3), 402-407.

ANEXOS

Nº 1: El instrumento aplicado

EL DISEÑO DE INVESTIGACIÓN

Una vez definido el tipo de estudio por realizar y establecidas las hipótesis o los lineamientos para la investigación (si no se tiene una hipótesis), el investigador debe concebir un *diseño de investigación*. Según su acepción de “esbozo o bosquejo”, es el plan o la concebida para responder a las de investigación, para realizar la que toda investigación científica (lo que constituye la de la investigación empírica). la manera práctica y de confrontar la hipótesis con la

En la literatura, encontramos clasificaciones. Nosotros adoptaremos la : investigación no experimental investigación experimental.

¿Qué es la investigación no experimental?

Es la que se *sin manipular deliberadamente* las En ella, observamos el tal y como se en su contexto natural; hay estímulos a los se expongan los sujetos estudio.

Se clasifica por alcance temporal, por el de momentos (puntos en el) en los cuales se los datos en: seccional longitudinal.

La primera analiza es el estado de una o variables en un momento , o cuál es la entre las variables en un en el tiempo, o parte de la del fenómeno ya producido, y de hallar las causas o que determinan su aparición. veces, en cambio, la se centra en estudiar evolucionan la/s variable/s o la relación entre En estas situaciones, el apropiado (bajo el enfoque experimental) es un estudio Puede ser a “corto” (un año o menos, un ciclo lectivo) o a “.....” plazo, menos frecuente (por , el estudio clásico de Terman niños geniales que duró 20).

¿Qué es la investigación experimental?

El término *experimento* tiene dos, una general y otra
La general se refiere a “..... una acción” y observar sus Este uso del término es coloquial; así hablamos de “experimentar” inventamos o adaptamos una de cocina y probamos el

El *sentido estricto*, más con el uso científico del, remite a una investigación la que se manipulan una o más variables (supuestas causas) para analizar su sobre una o más variables (supuestos), en una situación de control.

Nº 2: La clave de corrección

EL DISEÑO DE INVESTIGACIÓN

Una vez definido el tipo de estudio por realizar y establecidas las hipótesis o los lineamientos para la investigación (si no se tiene una hipótesis), el investigador debe concebir un *diseño de investigación*. Según su acepción de “esbozo o bosquejo”, es el plan o la **1) estrategia** concebida para responder a las **2) preguntas** de investigación, para realizar la **3) prueba** que toda investigación científica **4) supone** (lo que constituye la **5) esencia** de la investigación empírica). **6) Es** la manera práctica y **7) concreta** de confrontar la hipótesis con la **8) realidad**.

En la literatura, encontramos **9) diversas** clasificaciones. Nosotros adoptaremos la **10) siguiente**: investigación no experimental **11) e** investigación experimental.

¿Qué es la investigación no experimental?

Es la que se **12) realiza sin manipular deliberadamente** las **13) variables**. En ella, observamos el **14) fenómeno** tal y como se **15) presenta** en su contexto natural; **16) no** hay estímulos a los **17) que** se expongan los sujetos **18) del** estudio.

Se clasifica por **19) su** alcance temporal, por el **20) número** de momentos (puntos en el **21) tiempo**) en los cuales se **22) recolectan** los datos en: seccional **23) o** longitudinal.

La primera analiza **24) cuál** es el estado de una o **25) más** variables en un momento **26) dado**, o cuál es la **27) relación** entre las variables en un **28) punto** en el tiempo, o parte de la **29) observación** del fenómeno ya producido, y **30) trata** de hallar las causas o **31) factores** que determinan su aparición.

32) Otras veces, en cambio, la **33) investigación** se centra en estudiar **34) cómo** evolucionan la/s variable/s o la relación entre **35) ellas**. En estas situaciones, el **36) diseño** apropiado (bajo el enfoque **37) no** experimental) es un estudio **38) longitudinal**. Puede ser a “corto” **39) plazo** (un año o menos, **40) usualmente** un ciclo lectivo) o a “**41) largo**” plazo, menos frecuente (por **42) ejemplo**, el estudio clásico de Terman **43) sobre** niños geniales que duró **20 44) años**).

¿Qué es la investigación experimental?

El término *experimento* tiene dos **45) acepciones**, una general y otra **46) particular**. La general se refiere a “**47) ejecutar** una acción” y observar sus **48) consecuencias**. Este uso del término es **49) bastante** coloquial; así hablamos de “experimentar” **50) cuando** inventamos o adaptamos una **51) receta** de cocina y probamos el **52) resultado**.

El *sentido estricto*, más **53) acorde** con el uso científico del **54) término**, remite a una investigación **55) en** la que se manipulan **56) deliberadamente** una o más variables **57) independientes** (supuestas causas) para analizar su **58) incidencia** sobre una o más variables **59) dependientes** (supuestos **60) efectos**), en una situación de control.