

Evaluación de la comprensión del texto científico-académico en el posgrado mediante el procedimiento cloze.

Difabio de Anglat, Hilda y Álvarez, Guadalupe.

Cita:

Difabio de Anglat, Hilda y Álvarez, Guadalupe (2021). *Evaluación de la comprensión del texto científico-académico en el posgrado mediante el procedimiento cloze*. e-ISSN: 1667-4545 - Revista Evaluar, 21, 17-38.

Dirección estable: <https://www.aacademica.org/guadalupe.alvarez/130>

ARK: <https://n2t.net/ark:/13683/p0zR/ghb>



Esta obra está bajo una licencia de Creative Commons.
Para ver una copia de esta licencia, visite
<https://creativecommons.org/licenses/by-nc-nd/4.0/deed.es>.

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.



Evaluación de la comprensión del texto científico-académico en el posgrado mediante el procedimiento cloze

Assessment of Reading Comprehension of Scientific-Academic Text at the Postgraduate Level Using Cloze Procedure

Introducción
Método
Resultados
Discusión
Referencias

Hilda Difabio de Anglat *¹, Guadalupe Álvarez²

1 - Centro de Investigaciones Cuyo Dr. Abelardo Pithod, CONICET, Mendoza, Argentina.

2 - Instituto del Desarrollo Humano de la Universidad Nacional de General Sarmiento, CONICET, Buenos Aires, Argentina.

Recibido: 14/10/2020 Revisado: 25/11/2020 Aceptado: 15/12/2020

Resumen

El cloze es un procedimiento que reviste potencial para evaluar de forma parsimoniosa la comprensión del texto científico-académico, para el cual no se registran antecedentes de investigación respecto de la comprensión de textos en español en el nivel de posgrado. Por ello, los objetivos de este trabajo fueron: 1) identificar el nivel de dificultad e índice de discriminación de los ítems de un cloze que se empleó en un curso de posgrado (Facultad de Filosofía y Letras, Universidad Nacional de Cuyo, Argentina); 2) ponderar si la distribución desde ítems muy fáciles a muy difíciles se ajustaba a los porcentajes aconsejados; 3) describir los niveles funcionales de lectura de los participantes, y 4) estimar la incidencia, en la dificultad de los ítems, de los denominados factores intrínsecos o aspectos de la tarea. En un corpus conformado por 82 clozes, mediante análisis estadísticos, se hallaron buenas cualidades psicométricas para el instrumento y niveles diferenciales de dificultad de las palabras lexicales en función de los factores en estudio.

Palabras clave: *comprensión lectora, procedimiento cloze, propiedades psicométricas, nivel de posgrado, español*

Abstract

Cloze is a procedure with the potential of parsimoniously assessing the understanding of the scientific-academic text. There are no previous studies regarding the comprehension of Spanish texts at the postgraduate level. The aims of this study were: 1) to identify difficulty levels and discrimination index of the cloze items used in a postgraduate course (Facultad de Filosofía y Letras, Universidad Nacional de Cuyo, Argentina); 2) to analyse whether the distribution of the items ranging from very easy to very hard corresponds to percentages advised by statistical theory; 3) to describe the participants' functional reading levels and 4) to estimate the incidence of intrinsic factors or aspects of the task in items' difficulty. Eighty-two cloze tests were collected and statistically analysed. The analysis revealed acceptable psychometric qualities for the instrument and acceptable differential levels of difficulty in content words according to the factors under study.

Keywords: *reading comprehension, cloze procedure, psychometric properties, postgraduate level, Spanish*

* **Correspondencia a:** Hilda Difabio de Anglat. Primitivo de la Reta 522, 2° K, 5.500 Mendoza, Argentina. E-mail: ganglat@gmail.com

Introducción

La importancia de la comprensión de textos científico-académicos en los estudios de posgrado es indiscutible en sí misma y en su contribución al desarrollo de los procesos de producción textual (Parodi, 2001) por las “conexiones de tipo psicolingüístico y discursivo” (p. 59) que se verifican entre estos dos procesos y porque, en este nivel educativo, se lee para escribir. En consecuencia, en una instancia educativa orientada a promover la producción escrita (un taller virtual focalizado en familiarizarse con el género para componer un primer capítulo o apartado de la tesis), se pondera la competencia lectora de los cursantes mediante un cloze, procedimiento que suprime palabras de un texto en prosa, reemplazándolas por una línea de extensión constante.

Wilson Taylor (1953) desarrolló el test de cloze como medida de la legibilidad o comprensibilidad de los textos, aunque el procedimiento ha sido empleado en forma ininterrumpida como herramienta de investigación y de evaluación de la competencia lectora, e inclusive como dispositivo de enseñanza (p. ej., Sadeghi, 2008).

El término *cloze* parece ser un neologismo, pues no figura en los diccionarios (Cairney, 2018). Taylor (1953) señala que lo deriva de la palabra inglesa *closure* en tanto constituye una aplicación al lenguaje del principio de cierre de la psicología de la Gestalt o tendencia a percibir como completa una forma fragmentada, a completar un patrón una vez que se ha captado su significado esencial.

Para completar un cloze, el lector debe *cerrar* las percepciones truncadas mediante inferencias (p. ej., Chino-Vilca & Zegarra-Valdivia, 2019) a partir de las claves semánticas (entre ellas, la redundancia textual), sintácticas, discursivas, y las restricciones del lenguaje (Peña-Calvo, 1997), así como de su conocimiento previo y de su expectativa intuitiva (Difabio de Anglat, 2008).

Esto es, el procedimiento requiere que el lector active todos sus recursos cognitivos y lingüísticos a fin de sostener el proceso interactivo entre la información que aporta el texto (proceso de abajo hacia arriba) y la que él anticipa (proceso de arriba hacia abajo); gracias a estos procesos, el lector cumple el objetivo propio de toda tarea de lectura: “dotar de coherencia a lo leído” (León-Cascón, 2004, p. 109). Para ello, el cloze exige *reconstruir el texto* a un nivel de mayor control o monitoreo que la lectura normal (esto es, de un texto completo) ya que demanda atender a una gama amplia de señales adicionales. La comprensión de un cloze también reclama considerable motivación e interés para mantener la atención en los procesos referidos. Por ende, no solo parece contribuir a generar mayor aprehensión del significado, sino a estimular la reflexión metacognitiva sobre el desempeño o *comprensión metacognitiva* (León-Cascón, 2004). Por otra parte, dado que es relativamente resistente al tiempo de exposición, es apropiado para la autoadministración (Datta & Macdonald-Ross, 2002).

Sin embargo, aún hoy se investigan los problemas planteados en los estudios tempranos (Abraham & Chapelle, 1992; Brown, 1980, 1988; Jonz, 1990), que continúan siendo relevantes; entre ellos, su validez, qué evalúa exactamente (p. ej., Gellert & Elbro, 2013), los patrones de eliminación de términos (Kleijn, Pander-Maat, & Sanders, 2019), los métodos de calificación (Kleijn et al., 2019; O’Toole & King, 2010) y, en menor medida, la dificultad de los ítems (Trace, 2020; Trace, Brown, Janssen, & Kozhevnikova, 2017). La primera pregunta conforma un área de ininterrumpido debate, que enfrenta a quienes consideran que solo evalúa comprensión local (por ejemplo, Carlisle & Rice, 2004, en Trace et al., 2017) con aquellos que sostienen que puede ponderar la coherencia global, constituyéndose así en una medida válida y rápida de la competencia lectora

(Gellert & Elbro, 2013; Kleijn et al., 2019; Shah-nazari-Dorcheh, Roshan, & Hesabi, 2012; Trace, 2020).

Esto es, aunque el cloze no produzca pruebas perfectas de la competencia lectora general, reviste potencial para evaluar de forma parsimoniosa las habilidades superiores de lectura (el monitoreo de la comprensión, la anticipación o predicción, la inferencia, la elaboración o relación del texto con el conocimiento previo, la vinculación entre diferentes ideas expuestas en el texto), si las eliminaciones deliberadamente apelan a demandas inferenciales vinculadas con la estructura lógica subyacente al texto (Green, 2001). En otras palabras, se pueden generar lagunas de mayor interés (Gellert & Elbro, 2013) por la supresión de palabras de contenido (palabras léxicas) y conectores que se basan en las relaciones intertextuales.

Para la calificación de un texto cloze, se dispone de dos métodos: aceptar solo la palabra efectivamente omitida (*verbatim*), *puntuación de la respuesta exacta* (Brown, 2013), o computar también los sinónimos, *puntuación de la respuesta aceptable*, calificación semántica (Kleijn, 2018) o conceptual (O'Toole & King, 2011).

Hoy los investigadores (Brown, 2013; Kleijn, 2018; Kleijn et al., 2019; O'Toole & King, 2011) acuerdan en que la segunda es una puntuación más adecuada en términos de discriminación de los ítems (una cualidad importante para nuestra investigación) y, sobre todo, porque proporciona una ponderación más ajustada del desempeño del lector en cuanto la respuesta exacta enmascara la competencia que pueden reflejar reemplazos potencialmente lícitos (Kobayashi, 2002). Al respecto, el estudio meta-analítico de Watanabe y Koyama (2008) confirma la investigación seminal de Brown (1980), poniendo en evidencia que el método de puntuación *verbatim* es menos confiable y válido que el de la respuesta aceptable,

en particular en relación con las palabras léxicas.

La estrategia usual para la interpretación de los resultados consiste en transformar los porcentajes finales en niveles funcionales de lectura: independiente, instruccional y de frustración (p. ej., Condemarín & Milicic, 1988). El primero remite a un lector que comprende acabadamente el texto de modo autónomo. El nivel instruccional significa una aproximación global, con algunas dificultades en el reconocimiento de palabras y en el proceso inferencial. En el de frustración, la comprensión es deficiente porque el lector se apoya solo en el contexto inmediato para completar los blancos. Con relación al criterio para la adscripción de estos niveles, los trabajos en español (entre otros, Calderón-Ibañez & Quijano-Peñuela, 2010; Chino-Vilca & Zegarra-Valdivia, 2019) aplican la categorización porcentual de González-Moreyra (1998).

Con referencia a la investigación de la última década en dicho idioma, varios estudios emplean la técnica para evaluar la competencia lectora de estudiantes de nivel universitario, ya sea como instrumento único (p. ej., Calderón-Ibañez & Quijano-Peñuela, 2010; Márquez, Ibañez, & Pérez, 2011) o en combinación con otros (Chino-Vilca & Zegarra-Valdivia, 2019; Esquivel-Gámez, Martínez-Olvera, Córdoba del Valle, & Reyes-Gutiérrez, 2016; Tabullo et al., 2020). En general, estos estudios han mostrado que el nivel de comprensión de los alumnos se halla por debajo del deseado y que el desempeño en el cloze se relaciona con el rendimiento académico.

En cambio, hasta donde se ha podido indagar, el cloze no se utiliza con frecuencia en el posgrado; los dos trabajos que se identificaron (Costa, 2006; Sadeghi, 2008) lo aplican en portugués e inglés, respectivamente. De allí el interés por ponderar su potencialidad evaluativa en lectores expertos de español como primera lengua; en particular, se intenta analizar los factores que inciden

en la dificultad de los ítems. Este propósito registra trabajos tempranos (Abraham & Chapelle, 1992; Bachman, 1985; Brown, 1989; Jonz, 1990; Kobayashi, 2002) y más recientes desde la teoría de respuesta al ítem (p. ej., Trace, 2020; Trace et al., 2017), análisis inaplicable en el corpus que se analiza por el tamaño de la muestra, dado que “para emplear modelos TRI se requieren muestras grandes de sujetos ($n > 300$)” (Cortada de Kohan, 2004, p. 105).

Abraham y Chapelle (1992) estudian los *factores intrínsecos*, que definen como aspectos de contenido de la tarea factibles de medición con independencia de las características del lector. Las autoras consideran: palabras de contenido o funcionales y, para las primeras, cantidad de contexto requerido para completar los blancos, viabilidad de respuestas alternativas y número de ocurrencias del término en el texto.

A partir de las consideraciones previas, el presente trabajo se plantea cuatro objetivos: 1) identificar el nivel de dificultad e índice de discriminación de los ítems de un cloze que se emplea en un taller virtual de escritura en el posgrado; 2) ponderar si la distribución desde ítems muy fáciles a muy difíciles se ajusta a los porcentajes aconsejados; 3) describir los niveles funcionales de lectura de los participantes, y 4) estimar la incidencia, en la dificultad de los ítems, de los referidos factores intrínsecos, nuestro objetivo principal.

Metodología

Diseño y participantes

Se desarrolló un estudio empírico con metodología cuantitativa, transversal y de profundidad analítica en un corpus de textos cloze proveniente de los cursantes de Estrategias para la producción conceptual y escrita en el posgrado en Ciencias

Sociales y Humanas, taller en línea acreditado por la Secretaría de Posgrado, Facultad de Filosofía y Letras, Universidad Nacional de Cuyo (UNCuyo), Mendoza, Argentina. En este curso, que se implementa en la plataforma Moodle institucional, el cloze forma parte de un diagnóstico integral. A partir de la consigna proporcionada (ver Anexo 1), cada asistente lo completa de modo asincrónico durante la primera semana; luego, sube la resolución al espacio creado con este fin en la plataforma.

La muestra ($N = 82$) se conforma por 76 doctorandos, cinco maestrandos y una doctora en Psicología; 60 mujeres y 22 varones; todos hablantes nativos del español. Los estudiantes se distribuyen entre carreras de posgrado en Educación (el 39%), Letras (21%) y otras disciplinas (Filosofía, Historia, Geografía, Psicología, Sociología, Derecho y Arquitectura) de la UNCuyo pero también de otras universidades argentinas. Diez de los asistentes son extranjeros, todos procedentes de países hispanoamericanos.

La muestra se constituyó de modo progresivo entre 2015 y 2019. Su tamaño se debe a que este taller, en pro de alcanzar su objetivo último de favorecer la producción de un capítulo o apartado de la tesis, es de cupo limitado (hasta 12 estudiantes). Dado que los datos provienen de las implementaciones del cloze en sucesivas versiones del curso referido (en total, siete), por un lado, se recurre al test no paramétrico de Kruskal-Wallis para establecer si se verifican diferencias de rendimiento significativas entre los grupos; por el otro, a la prueba de chi cuadrado a fin de confrontar los niveles funcionales de lectura entre cursantes con formación en Letras ($n = 17$) y sin dicha formación ($n = 65$). Los resultados ($\chi^2 = 3.412$, $gl = 5$, $p = .756$; $\chi^2 = 3.144$, $gl = 2$, $p = .208$, respectivamente) ponen de manifiesto que es posible reunirlos en una sola muestra.

Instrumento

El cloze aplicado (ver Anexo 1) sintetiza el texto “Principios para un final: reflexiones en torno a la escritura de la tesis” (Álvarez, 2011), al que se le agregó como cierre parte de un párrafo de Cubo de Severino y Bosio (2011), capítulo del libro de cabecera del taller. En consecuencia, recurre a un texto ya publicado, en lugar de a uno elaborado *ad hoc*, porque la evidencia empírica (p. ej., O’Toole & King, 2011) ha demostrado que emplear un texto auténtico incrementa la validez de la prueba. Además, conforma un ejemplar representativo del tipo de lectura que los estudiantes efectuarán durante el curso y no depende de conocimiento especializado (palabras *impredecibles*) en cuanto se garantizó al menos una ocurrencia en el texto de los términos técnicos (*situación comunicativa, puesta en discurso, revisión, edición*).

El cloze consta de 484 palabras y 79 blancos, los que se establecieron mediante la combinación de dos criterios o estrategia híbrida (método semejante al de Kleijn et al., 2019): *razón fija*, que consiste en la eliminación del quinto vocablo y sus múltiplos, y *racional*, el cual permite controlar las palabras omitidas y, por esta vía, maximizar el empleo de información fuera del párrafo que contiene el blanco (Abraham & Chapelle, 1992; Bachman, 1985; Gellert & Elbro, 2013). En este sentido, se seleccionó un porcentaje ínfimo (16%) de términos que se podían completar con el solo conocimiento gramatical (palabras funcionales), y el 60% de las palabras léxicas exigía atender a claves entre párrafos o anticipar términos que no aparecen en el texto. El porcentaje total de omisión (también del 16%) se ubicó entre los estándares de la razón fija (20%) y del criterio racional (10%; Kleijn, 2018). La primera oración y la última se reprodujeron completas.

En la calificación se atribuyó un punto a

cada laguna correctamente resuelta (*verbatim* o un sinónimo). El criterio para la determinación de las alternativas correctas fue su adecuación global, esto es, la respuesta debía cumplir “los requisitos de todo el contexto del discurso en el que aparece” (Oller & Jonz, 1994, en Kleijn et al., 2019, p. 8). El total posible fue de 79 puntos, puntuación que se transformó en porcentaje, dado que los niveles funcionales de lectura se expresan porcentualmente (Condemarín & Milicic, 1988; Difabio de Anglat, 2008).

Análisis de datos

Se llevaron a cabo los siguientes análisis. En primer lugar, se elaboró el espacio de alternativas correctas (ver Anexo 2), una tarea importante para una investigación basada en la puntuación de la respuesta aceptable. Por ello, se efectuó mediante acuerdo inter-jueces a través del siguiente procedimiento: la primera autora confeccionó una matriz inicial de todas las respuestas (6.478), distinguiendo entre *verbatim*, alternativa/s correcta/s y error/es; la segunda autora evaluó la codificación y señaló algunos desacuerdos (de alrededor del 15%); finalmente, una especialista en lingüística del español revisó el análisis, lo que permitió resolver las disidencias.

En segundo lugar, se determinó el nivel de dificultad (ND) de cada ítem; como es práctica frecuente (Abraham & Chapelle, 1992; Brown, 1989, 2013; Kobayashi, 2002; Angeli Dos Santos, Primi, Taxa, & Vendramini, 2002), se empleó un método sencillo, aplicado en un trabajo previo (Difabio de Anglat, 2008), que consiste en dividir, en cada ítem, la frecuencia de respuestas correctas por el tamaño de la muestra. En las investigaciones sobre el cloze (p. ej., Abraham & Chapelle, 1992; Brown, 1989) este porcentaje se denomina ND y no nivel de facilidad de los ítems.

Para la interpretación de los resultados, se utilizó la categorización de Backhoff-Escudero, Larrazo-Reyna y Rosas-Morales (2000) en ítems muy fáciles (.87 a 1), fáciles (.74 a .86), moderados (.53 a .73), difíciles (.33 a .52) y muy difíciles (.00 a .32).

En tercer lugar, se estableció el índice de discriminación (ID) de cada ítem. Se aplicó también el procedimiento clásico (Brown, 1989; Angeli Dos Santos et al., 2002) que consiste en restar, en cada ítem, los niveles de dificultad hallados en los grupos extremos (superior, GS, e inferior, GI); estos grupos se conformaron según la puntuación total por el 27% superior y el 27% inferior, 22 cursantes en cada uno.

En cuarto lugar, se calculó el índice de dificultad del *cloze bien adaptado* —*well-tailored cloze* (Brown, 1988, 2013)—. Un *cloze* bien adaptado es aquel que retiene los ítems que manifiestan índices de discriminación $\geq .30$ (Lozano-Fernández & de la Fuente-Solana, 2015). En estos índices se distinguió entre buena potencialidad discriminativa (entre .30 y .39) y muy buena (.40 o más). No obstante, también se conservaron algunos ítems con discriminación marginal (entre .20 y .29).

En quinto lugar, las puntuaciones totales del *cloze* adaptado se transformaron en niveles funcionales de lectura o categorías de competencia lectora. Se adecuó la categorización de Gonzá-

lez-Moreyra (1998) para dicha transformación: *independiente*, puntajes de 75% o más, subcategorizado en tres niveles, excelente (100% - 90%), muy bueno (89% - 80%) y bueno (75% - 79%); *instruccional*, entre 59% y 74%, y *de frustración*, 58% o menos, subcategorizado en de dificultad (58% - 45%) y deficiente (44% o menos).

En sexto lugar, se estimó la incidencia en ND de los factores intrínsecos; se efectuaron cuatro análisis: 1) palabras léxicas vs funcionales, 2) cantidad de contexto requerido para completar el blanco, 3) número de respuestas alternativas y 4) otras ocurrencias del ítem en el texto. Los tres últimos análisis se restringieron a los términos léxicos porque la investigación ha mostrado (Abraham & Chapelle, 1992) que solo se justifican para este tipo de palabras.

De acuerdo con la investigación previa (Abraham & Chapelle, 1992; Bachman, 1985; Jonz, 1990; Kobayashi, 2002; Trace, 2020; Trace et al., 2017), en el presente trabajo se postula que tienden a ser más fáciles los ítems funcionales que los lexicales, los que requieren menor cantidad de contexto para restaurar las palabras suprimidas, los que registran alternativas conceptualmente aceptables y los que manifiestan otras ocurrencias en el texto.

A fin de determinar la cantidad de contexto requerido para completar cada blanco, se adaptó la categorización de la investigación clásica en

Tabla 1

Ejemplificación de las categorías sobre cantidad de contexto requerido para completar los blancos.

Categorías	Ejemplo/s
1. En la oración	Las operaciones de edición podrían describirse como <i>operaciones...</i>
2. En oraciones consecutivas	La revisión comprende dos subprocesos: <i>evaluación</i> y edición. En la evaluación se...
3. En el párrafo	La escritura, la presentación y la defensa de la tesis son las etapas conclusivas de los estudios de posgrado (...). Así, en el nivel de <i>posgrado</i> , el estudiante...
4. En el texto	La planificación es una evaluación inicial de la situación <i>comunicativa</i> (...).
5. Fuera del texto	El término “situación comunicativa” aparece completo al final del párrafo siguiente. <i>estrategias, elementos</i>

el tema (Abraham & Chapelle, 1992; Bachman, 1985; Jonz, 1990): 1) en la oración; 2) en oraciones consecutivas; 3) en el párrafo; 4) en el texto, y 5) fuera del texto ($n = 28$, el 41% de los blancos). Estas categorías se ejemplifican en la Tabla 1.

Para el análisis de esta variable, como en Abraham y Chapelle (1992), se aplicó un estadígrafo de correlación (ρ de Spearman por el tamaño de la muestra) entre las cinco categorías y el nivel de dificultad de los ítems expresado en porcentaje de respuestas correctas.

Dado que no se dispone de estudios previos sobre la incidencia del número de respuestas alternativas porque tanto Abraham y Chapelle (1992) como Kobayashi (2002) únicamente distinguen entre *una respuesta / más de una respuesta*, se computó dicho número a partir del corpus propio.

Finalmente, en relación con otras ocurrencias de las palabras de contenido en el texto, de los 58 blancos léxicos en el cloze aplicado aparecen 31 (el 53%) como término exacto o desde la familia de palabras en tres casos ($n^\circ 38$, *producir*, $n^\circ 56$, *elaborado* y $n^\circ 64$, *solucionar*), de los cuales únicamente cinco lo hacen más de dos veces. Luego, no resultó apropiada la categorización de Kobayashi (2002) en tres niveles: 1) la palabra no aparece en otro lugar del texto; 2) aparece entre dos y cuatro veces y 3) aparece cinco veces o más. De allí que, como en Abraham y Chapelle (1992), se distinguieron dos grupos según se verificaban o no ocurrencias en el texto (31 vs 27 palabras).

Los procedimientos analíticos empleados fueron: prueba de dependencia / independencia entre distribuciones empíricas (chi cuadrado o V de Cramer si el porcentaje de casillas con un recuento esperado menor que cinco superara el 20%) a fin de analizar la distribución de los niveles funcionales de lectura; diferencia entre muestras independientes y emparejadas (t de Student) en el caso de las variables clase de palabra (lexicales o

funcionales), número de respuestas alternativas y otras ocurrencias del término en el texto, y coeficiente de correlación (ρ de Spearman) respecto de la cantidad de contexto requerido para completar cada blanco.

Los análisis estadísticos se llevaron a cabo mediante el programa SPSS 25.0 (IBM Corporation, 2017).

Resultados

Antes de abordar los resultados referidos a la incidencia de los factores intrínsecos en la dificultad de los ítems, objetivo principal de este trabajo, cabe señalar que: a) los blancos sin responder ($n = 183$) representaron el 2.8%, un porcentaje similar al que permiten Kleijn et al. (2019) para retener una resolución del cloze en su muestra, en tanto lagunas en las que es lícito asumir que el lector no encuentra una respuesta porque están dispersas a lo largo de la prueba; b) la confiabilidad del cloze aplicado (alfa de Cronbach) es de .85, un buen índice que habilita ulteriores análisis ya que, obviamente, los resultados de un estudio son confiables en la medida en que lo es el instrumento en el que se basan; c) el rango de las correlaciones ítem-total (coeficiente de homogeneidad) se distribuyó entre .05 (ítem $n^\circ 39$, *como*) y .59 ($n^\circ 43$, *Puesta*), con una media de .31.; d) ocho ítems no alcanzaron .20, el mínimo requerido de homogeneidad (Kline, 1998); sin embargo, se retuvieron dado que su supresión incrementaba el alfa en solo un centésimo (.86). Al respecto, Kobayashi (2002) trabaja con una distribución cercana: entre .08 y .53, con una media de .34, en estudiantes universitarios japoneses que completan textos cloze en inglés.

En la Tabla 2, que se divide en dos para facilitar la lectura, se presentan los resultados del análisis del nivel de dificultad e índice de discri-

Tabla 2

Nivel de dificultad e índice de discriminación de los ítems del cloze.

Ítem	ND	GS	GI	ID	Ítem	ND	GS	GI	ID
1	.95	.95	.95	0	21	.74	.95	.64	.31
2	.72	.82	.50	.32	22	.79	.95	.55	.40
3	.79	.95	.68	.27	23	.73	.95	.50	.45
4	.40	.59	.27	.32	24	.72	.86	.55	.31
5	.98	1	.95	.05	25	.91	1	.77	.23
6	.55	.73	.36	.37	26	.41	.64	.27	.37
7	.48	.64	.32	.32	27	.49	.68	.36	.32
8	.54	.77	.32	.45	28	.73	.91	.55	.36
9	.78	.91	.59	.32	29	.49	.82	.14	.68
10	.90	1	.77	.23	30	.80	.95	.68	.27
11	.73	.95	.50	.45	31	.93	1	.86	.14
12	.73	.91	.55	.36	32	.90	1	.77	.23
13	.94	.95	.95	0	33	.83	1	.68	.32
14	.32	.45	.14	.31	34	.26	.50	.09	.41
15	.48	.68	.36	.32	35	.54	.68	.36	.32
16	.80	1	.50	.50	36	.79	1	.68	.32
17	.90	1	.77	.23	37	.95	1	.95	.05
18	.96	1	.91	.09	38	.87	1	.68	.32
19	.82	.95	.64	.31	39	.37	.55	.32	.23
20	.84	1	.68	.32	40	.69	.82	.45	.37

Ítem	ND	GS	GI	ID	Ítem	ND	GS	GI	ID
41	.96	.95	.95	0	61	.32	.68	.04	.64
42	.80	1	.68	.32	62	.54	.82	.45	.37
43	.57	.95	.05	.91	63	.44	.64	.23	.41
44	.82	1	.64	.36	64	.71	.91	.59	.32
45	.18	.45	0	.45	65	.90	.95	.86	.09
46	.35	.68	.18	.50	66	.71	.91	.45	.46
47	.54	.73	.36	.37	67	.84	1	.55	.45
48	.98	.95	.95	0	68	.80	1	.68	.32
49	.73	.86	.50	.36	69	.61	.77	.41	.36
50	.89	1	.68	.32	70	.90	1	.77	.23
51	.40	.64	.18	.46	71	.57	.73	.41	.32
52	.60	.68	.36	.32	72	.77	.91	.55	.36
53	.46	.82	.14	.68	73	.73	.77	.45	.32
54	.78	1	.50	.50	74	.96	1	.91	.09
55	.72	.95	.45	.50	75	.46	.68	.27	.41
56	.67	.77	.55	.22	76	.61	.73	.27	.46
57	.37	.55	.23	.32	77	.73	.91	.50	.41
58	.90	1	.77	.23	78	.24	.50	.14	.36
59	.62	.73	.27	.46	79	.60	.73	.41	.32
60	.80	.95	.64	.31					

Nota. ND = Nivel de dificultad; GS = Grupo superior; GI = Grupo inferior; ID = Índice de discriminación.

minación de los ítems.

Diez ítems no alcanzaron el mínimo aceptable de discriminación porque los respondieron correctamente la mayoría de los estudiantes (en-

tre dos y seis errores por ítem); ellos son: n° 1 (*ordenamiento*), n° 5 (*tesis*), n° 13 (*investigación*), n° 18 (*como*), n° 31 (*mutuamente*), n° 37 (*para*), n° 41 (*Es*), n° 48 (*cabo*), n° 65 (*nivel*) y n° 74

(*problemas*). En consecuencia, el instrumento se redujo a 69 ítems, de los cuales 58 corresponden a palabras léxicas y 11 a palabras funcionales, cantidad que mantiene la relación proporcional inicialmente intentada (84% vs 16%).

También 10 son los ítems que evidenciaron discriminación marginal (destacados en letra itálica en la Tabla 2), el 14.5% del instrumento, de los cuales seis corresponden a ítems muy fáciles: n° 10 (*producción*), n° 17 (*aspectos*), n° 25 (*resolución*), n° 32 (*proceso*), n° 58 (*hasta*) y n° 70 (*sobre*). Los restantes se distribuyeron entre discriminación buena —35 ítems (el 50.7%)— y muy buena —24 ítems (el 34.8%)—; cuatro manifiestan excelente potencialidad discriminativa (superior a .60) y remiten a términos técnicos: n°

29 (*discurso*), n° 43 (*Puesta*), n° 53 (*revisión*) y n° 61 (*tarea* de escritura). El ítem n° 43 alcanzó discriminación prácticamente óptima (.91); recordemos que también manifiesta el índice más alto de correlación ítem-test.

En relación con el nivel de dificultad de los ítems, mediante la categorización de Backhoff-Escudero et al. (2000) desde ítems muy fáciles a muy difíciles, se hallaron los resultados que se presentan en la Tabla 3.

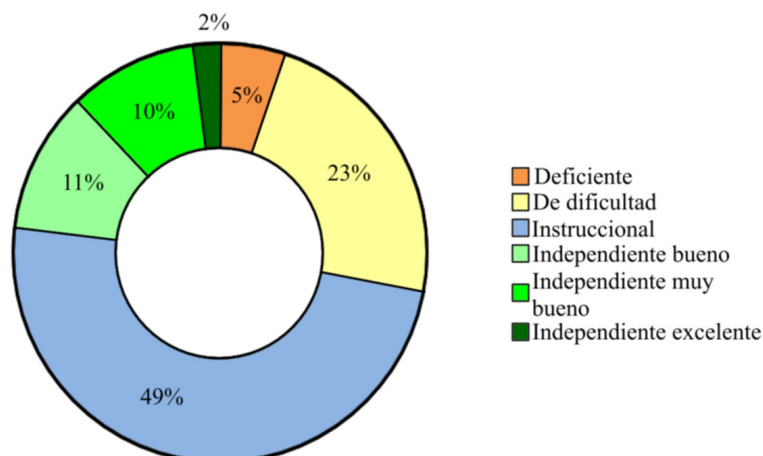
El índice medio de dificultad del cloze es .65, levemente superior al rango sugerido por la teoría estadística (entre .50 y .60; Brown, 1999).

El blanco más difícil es el n° 45 (*considerando*); solo 15 cursantes lo respondieron correctamente, con el gerundio esperado o una alternati-

Tabla 3

Categorización de los niveles de dificultad de los ítems del cloze.

Nivel de dificultad	Categorización de Backhoff-Escudero et al. (2000)	% aproximado ideal	% en este instrumento	Ítems N°
Muy fácil	.87 a 1	10%	10%	8 ítems: 10 (producción), 17 (aspectos), 25 (resolución), 32 (proceso), 38 (producir), 50 (por), 58 (hasta) y 70 (sobre)
Fácil	.74 a .86	20%	25%	17 ítems: 3 (escritura), 9 (estrategias), 16 (rol), 19 (dicho), 20 (concebirse), 21 (parte), 22 (cuya), 30 (se), 33 (implica), 36 (tiempo), 42 (etapa), 44 (expresar), 54 (evaluación), 60 (ajuste), 67 (operaciones), 68 (elementos) y 72 (procesos)
Moderado	.53 a .73	40%	38%	26 ítems: 2 (procesos), 6 (su), 8 (exige), 11 (aquellas, las), 12 (textos), 23 (elaboración), 24 (sentido), 28 (escrita), 35 (texto), 40 (lista), 43 (Puesta), 47 (detallado), 49 (conocer), 52 (requerido), 55 (verifica), 56 (elaborado), 59 (corrobora), 62 (adecuación), 64 (solucionar), 66 (edición), 69 (inserción), 71 (hacer), 73 (partida), 76 (además), 77 (reflexión) y 79 (razones)
Difícil	.33 a .52	20%	19%	13 ítems: 4 (resulta), 7 (dominio), 15 (posgrado), 26 (realizar), 27 (involucrados), 29 (discurso), 39 (como), 46 (escritura), 51 (lenguaje), 53 (revisión), 57 (su), 63 (ortográficas) y 75 (esta)
Muy difícil	.00 a .32	10%	7%	5 ítems: 14 (Así), 34 (comunicativa), 45 (considerando), 61 (tarea) y 78 (académicos)



va aceptable (por ejemplo, *utilizando*). El resto de los blancos de esta categoría remite a un adverbio (n° 14, *Así*), dos adjetivos (n° 34, *comunicativa*, y n° 78, *académicos*) y un sustantivo (n° 61, *tarea de escritura*); todos términos de contenido.

Los niveles funcionales de lectura hallados en la muestra pueden verse en el Gráfico 1.

En el nivel de frustración se ubicó el 28% de la muestra; por carrera, el porcentaje menor de cursantes en este nivel correspondió a los historiadores (4%), el siguiente (9%), a los maestrandos o doctorandos en Letras, Derecho, Filosofía y Psicología. El resto de los estudiantes se distribuyó de la siguiente manera: el 49% en el nivel instruccional y el 23% en el nivel independiente. Ninguna de estas diferencias por carrera alcanzó significatividad estadística (V de Cramer = .294, $p = .292$).

En relación con la incidencia de los factores intrínsecos en la dificultad de los ítems, en el primer análisis (palabras léxicas vs funcionales), la prueba t de Student (palabras funcionales: $X = 69.74$, $DE = 16.54$; palabras léxicas: $X = 63.95$, $DE = 13.61$) evidencia diferencias estadísticamente significativas ($t = 3.704$, $gl = 81$, $p < .001$), siendo superior el valor de la media en el caso de las palabras funcionales. Ya que la divergencia de medias en términos absolutos no es tan elocuente,

se realizó un análisis adicional. Se calcularon los niveles de lectura en la puntuación total en cada clase de palabras en las tres categorías principales: independiente, instruccional y de frustración. Luego, en una tabla cruzada (ver Tabla 4), se identificó el número de respuestas en las que: 1) se obtuvo el mismo nivel de rendimiento para ambas clases (43, el 53%); 2) fue superior el rendimiento en las palabras funcionales (28, el 34%) y 3) fue superior el desempeño en los términos léxicos (11, el 13%). Los resultados ($\chi^2 = 16.753$, $gl = 4$, $p < .01$) muestran que la diferencia entre 2) y 3) alcanza significatividad estadística.

Respecto de la cantidad de contexto requerido para completar cada blanco, el índice ($\rho = -.31$, $p < .05$) alcanzó significatividad estadística (es negativo por el modo de adscribir los números a las categorías de contexto), pero significa correlación débil (Hernández-Sampieri, Fernández-Collado, & Baptista-Lucio, 2014). Esto es, si bien se halló la progresión esperada en las cuatro primeras categorías (la media de ND de los ítems en la oración es .80; en oraciones consecutivas, .77; en el párrafo, .58; en el texto, .56), por un lado, son cercanos los índices de los pares primeros y de los segundos y, por otro, los ítems que se ubican en la categoría *fuera del texto*, en promedio, resultaron de dificultad moderada (.61).

Tabla 4

Niveles de lectura de las palabras léxicas y funcionales.

		Funcionales			Total
		De frustración	Instruccional	Independiente	
Léxicas	De frustración	10	10	4	24
	Instruccional	4	23	14	41
	Independiente	1	6	10	17
Total		15	39	28	82

A fin de analizar la incidencia del número de respuestas alternativas, esta variable se categorizó desde la mediana (rango = 0 a 15; mediana = 4) en dos grupos: 1) sin alternativa (términos técnicos) y hasta tres; y 2) cuatro o más alternativas. En el primer grupo, se ubicaron 24 palabras; en el segundo, las restantes 34.

La prueba *t* de Student manifestó que las diferencias ($X = .54$, $DE = .22$; $X = .70$, $DE = .13$) alcanzan significatividad estadística ($t = -3.514$, $gl = 56$, $p < .01$). En efecto, solo dos términos del grupo con más alternativas se asignaron a difíciles (n° 26, *realizar*, y n° 51, *lenguaje*); el resto se distribuyó desde ND moderado a muy fácil. Y viceversa: en el subgrupo sin / con menos alternativas se ubicaron seis ítems moderados (25%), ocho de los diez difíciles (n° 4, *resulta*; n° 7, *dominio*; n° 15, *posgrado*; n° 27, *involucrados*; n° 29, *discurso*; n° 46, plan de *escritura*; n° 53, *revisión*; y n° 63, *ortográficas*) y los cinco blancos muy difíciles.

Finalmente, en relación con otras ocurrencias de las palabras de contenido en el texto, las diferencias ($X = .66$, $DE = .19$; $X = .61$, $DE = .17$) no resultaron estadísticamente significativas ($t = 1.503$, $gl = 56$, $p = .297$).

En consecuencia, en el grupo con ocurrencias interesó analizar la asociación de esta variable con la cantidad de contexto requerido para completar cada blanco. Para ello, se diferenciaron dos niveles según se hallaren: 1) dos o más ocurrencias o una muy cercana (22 blancos) y 2) una lejana (9 blancos). La combinación de estos

criterios mostró que las diferencias ($X = .75$, $DE = .14$; $X = .46$, $DE = .16$) fueron estadísticamente significativas ($t = 5.136$, $gl = 29$, $p < .001$). En efecto, mientras que en el primer subgrupo se ubicó un único ítem difícil (n° 46, plan de *escritura*) y el resto se distribuyó desde ND moderado a muy fácil, en el segundo se posicionaron un blanco fácil (n° 42, *etapa*), uno moderado (n° 43, *Puesta*), cuatro difíciles (n° 15, *posgrado*, n° 29, *discurso*, n° 51, *lenguaje*, y n° 53, *revisión*) y tres muy difíciles (n° 14, *Así*, n° 34, *comunicativa*, y n° 61, *tarea* de escritura).

Discusión

El principal objetivo del presente trabajo consistió en identificar los factores intrínsecos al cloze que afectan el nivel diferencial de dificultad de los ítems: palabras de contenido o funcionales y, en el caso de las primeras, cantidad de contexto requerido para completar los blancos, número de respuestas alternativas y de otras ocurrencias del término en el texto. La relevancia de este análisis, a nuestro juicio, reside en que —a excepción de una investigación temprana y preliminar (Difabio de Anglat, 2008)— solo se dispone de este tipo de análisis respecto de textos en inglés y en portugués.

El instrumento aplicado resultó más sencillo que el ideal en cuanto aumenta el porcentaje de ítems fáciles y disminuye el correspondiente a muy difíciles. Sin embargo, a pesar del método

de calificación (palabra aceptable), en el nivel de frustración se ubicó una proporción alta de participantes, si se tiene en cuenta que son graduados universitarios que leen un cloze con un nivel promedio adecuado de dificultad (el cloze administrado incluso supera levemente dicho nivel promedio). Si bien podría pensarse que el tema del texto explica este resultado, también los doctorandos en Letras se posicionan en este nivel en porcentaje similar a los cursantes con formación en Derecho, Filosofía y Psicología.

Con relación a la incidencia de los factores intrínsecos en el nivel de dificultad (ND) de los ítems, los resultados parecen confirmar tres de los supuestos iniciales: tienden a ser más fáciles los ítems funcionales que los lexicales, los que requieren menor cantidad de contexto para restaurar las palabras suprimidas y los que registran alternativas conceptualmente aceptables. Sin embargo, no se verificó que el ND de los ítems fuera afectado por el número de ocurrencias de los términos en el texto.

Respecto del primer supuesto, en consonancia con trabajos en lengua inglesa (Abraham & Chapelle, 1992; Brown, 1989; Kobayashi, 2002; Trace et al., 2017), portuguesa (p. ej., Angeli Dos Santos et al., 2002; De Oliveira, 2011) y española (Difabio de Anglat, 2008), se confirma la dificultad superior de recuperación de las palabras léxicas en comparación con las funcionales (por ello, en pro de incrementar la entropía, el instrumento *ex profeso* incluyó un porcentaje considerablemente mayor de las primeras: 84%). Mientras que las palabras léxicas pertenecen a clases abiertas y contienen información semántica, las funcionales corresponden a clases sintácticas cerradas con un número restringido de ítems que son conocidos por prácticamente todos los hablantes. Es evidente, entonces, la probabilidad de su predicción acertada. Inclusive en las palabras funcionales dos estudiantes (sorprendentemente

doctorandos con formación en Educación, no en Letras) lograron resolución perfecta, desempeño que no se halló respecto de las léxicas. En suma, las primeras resultan más sencillas que las segundas, aun cuando tres blancos se ubicaron entre los ítems difíciles (n° 39, *como*, n° 57, *su*, y n° 75, *esta*); los dos últimos manifiestan la dificultad de la referencia anafórica, exigida por la precisión esperada en un texto académico.

Asimismo, el blanco más difícil corresponde al gerundio (*considerando*), una clase de palabra de contenido. La mayor dificultad del gerundio podría explicarse por su carácter de unión sin marca alguna de dos palabras (*que considera*); por otra parte, se trata de un gerundio adjunto o circunstancial, función que establece una relación “menos directa con el verbo principal” que otras funciones del gerundio (Real Academia Española, 2010, p. 513).

Respecto de la cantidad de contexto requerido para completar cada blanco, el índice de correlación en la muestra en estudio es cercano al que informan Abraham y Chapelle (1992) y levemente más alto que el de Kobayashi (2002); los tres coeficientes indican correlación débil. En este sentido, si bien algunos ítems que pueden completarse con información colindante (esto es, con información del contexto inmediato) resultan fáciles (p. ej., n° 25, *resolución*), otros también lo son (p. ej., n° 68, *elementos*) a pesar de requerir un término que no aparece en el texto. Este resultado podría asociarse con la frecuencia de uso de los vocablos desde el Corpus de Referencia del Español Actual —CREA— en cuanto, para cada término, presenta la frecuencia de uso absoluta y normalizada, un análisis de interés para un estudio futuro. Esta variable parece explicar la recuperación prácticamente perfecta, y consiguiente falta de discriminación, de expresiones muy familiares para graduados universitarios (por ejemplo, llevar a *cabo*).

En el caso de los blancos que mostraron rangos variables de ND desde moderado a muy difícil, cabe la observación de Kobayashi (2002) sobre la vinculación compleja de los términos que primordialmente transmiten significados con varias partes del texto; ello dificulta la localización fehaciente de las claves contextuales con las que dichas palabras se relacionan. Al respecto, solo Jonz (1990) proporciona criterios de identificación específicos, aunque priman pautas relativas a palabras funcionales —conjunciones, preposiciones, pronombres, marcadores del discurso, entre otras— (ver Jonz, 1990, *Appendix A*, pp. 76-77). De hecho, es el propio lector quien puede *brindar certeza* sobre qué información ha empleado para restaurar la palabra omitida. En un trabajo futuro, entonces, convendrá aplicar el cloze en entrevista (en línea, en nuestro caso, esto es, en el taller virtual de escritura que implementamos desde 2015, en el cual se administra el texto cloze como parte del diagnóstico) con el objeto de que el participante indique por qué eligió una respuesta en particular.

En cambio, sí es clara la relación entre número de respuestas alternativas y ND: en el subgrupo de más alternativas se ubica el 75% del total de ítems lexicales moderados y el 79% de los fáciles, mientras que en el subgrupo con pocas alternativas o una sola (términos técnicos) se halla el 80% de los difíciles y la totalidad de los muy difíciles. También Abraham y Chapelle (1992) encuentran una diferencia significativa de medias a favor de los blancos que admiten más de una respuesta. Por el contrario, Kobayashi (2002) informa medias cercanas pero en relación inversa (.36 vs .44), resultado que explica en función de la mayor demanda cognitiva que podrían exigir los términos que suponen un amplio rango de posibles respuestas. Si bien aparece como una argumentación plausible en el caso de la competencia semántica en una lengua extranjera (estudiantes

universitarios japoneses que completan textos cloze en inglés), para hablantes nativos instruidos (maestrandos y doctorandos) es razonable concluir que los blancos cuya resolución *verbatim* refiere a palabras ricas en sinónimos son más fáciles de completar.

A diferencia de los resultados de Kobayashi (2002), de Brown (1989) y de Trace et al. (2017), pero en consonancia con los de Abraham y Chapelle (1992), el nivel de dificultad de los ítems no es afectado por el número de ocurrencias de los términos en el texto. El primer autor encuentra una correlación débil que fundamenta en la asociación estrecha entre el tema del cloze y las palabras léxicas que se repiten. En Brown (1989) la combinación que alcanza la correlación múltiple mayor, con puntuación de la respuesta exacta, está conformada por cuatro variables: palabra de contenido o de función, frecuencia con la cual la respuesta correcta aparece en el cloze, más dos variables relativas a la comprensibilidad de un texto (número de caracteres de la palabra y de sílabas de la oración que contiene el blanco). Trace et al. (2017) informan que la cantidad de ocurrencias, conjuntamente con otras dos variables de la comprensibilidad (número de oraciones en un pasaje y de sílabas por palabra), explican el 24% de la varianza total en la dificultad de los ítems.

En nuestro caso, identificamos una asociación estadística entre ocurrencia y cantidad de contexto requerido para completar el blanco: los ítems con una ocurrencia lejana evidenciaron mayor dificultad que aquellos con más ocurrencias o una cercana. Esto es, como muestra la investigación previa (Abraham & Chapelle, 1992; Brown, 1989; Kobayashi, 2002; Trace, 2020; Trace et al., 2017) aunque con soluciones disímiles, el nivel diferencial de dificultad de las palabras de contenido no admite una explicación simple sino que exige el análisis conjunto de factores. Ya que transmiten significados y cada significa-

do individual opera en la conformación de la red semántica del texto (Kobayashi, 2002) y que obviamente dependen de la competencia del lector para aprovechar tanto las claves textuales como su conocimiento general, requieren una habilidad integradora de nivel superior. Al respecto, son los estudiantes con menor competencia lectora (esto es, aquellos que se posicionaron en el nivel de frustración en el cloze) quienes solo consideran el contexto inmediato para completar los blancos (el ejemplo paradigmático es el término técnico *situación “comunicativa”*, ítem cuya dificultad estriba en la necesidad de atender a la información entre párrafos).

En síntesis, el análisis efectuado parece apoyar el postulado de varios autores (Abraham & Chapelle, 1992; Brown, 2013; Gellert & Elbro, 2013; Shahnazari-Dorcheh et al., 2012; Trace, 2020, entre otros) sobre la validez del cloze (en dependencia de su método de construcción y calificación) para evaluar comprensión global, dado que la puntuación que obtiene el lector puede ser entendida como un indicador de su competencia para emplear las claves textuales. En este sentido, en la presente investigación la dificultad de los blancos referidos a lenguaje especializado varía de modo consistente según la *brecha* entre los ítems (Trace, 2020) —cantidad de contexto requerido para completarlos—, la que exige que se establezca la conexión entre proposiciones separadas en el texto (Abraham & Chapelle, 1992; Gellert & Elbro, 2013).

Por último, cabe enfatizar que la principal limitación de este estudio es el tamaño de la muestra que no posibilita análisis estadísticos más potentes; su incremento en un número considerable permitiría, por ejemplo, como en Brown (1989) y en Trace et al. (2017), un análisis de regresión a fin de determinar qué combinación de factores explica mejor la varianza de la dificultad de los ítems.

Referencias

- Abraham, R., & Chapelle, C. (1992). The meaning of cloze test scores: An item difficulty perspective. *The Modern Language Journal*, 76(4), 468-479. doi: 10.1111/j.1540-4781.1992.tb05394.x
- Álvarez, G. (2011). Principios para un final: Reflexiones en torno a la escritura de la tesis de Maestría. *Monográfico de la Maestría en Educación* (pp. 43-53). Medellín, Colombia: Universidad Pontificia Bolivariana.
- Angeli Dos Santos, A. A., Primi, R., Taxa, F. de O. S., & Vendramini, C. M. M. (2002). O teste de Cloze na avaliação da compreensão em leitura. *Psicologia: Reflexão e Crítica*, 15(3), 549-560. doi: 10.1590/S0102-79722002000300009
- Bachman, L. (1985). Performance on Cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly*, 19(3), 535-556. doi: 10.2307/3586277
- Backhoff-Escudero, E., Larrazolo-Reyna, N., & Rosas-Morales, M. (2000). Nivel de dificultad y poder de discriminación del Examen de Habilidades y Conocimientos Básicos (EXHCOBA). *Revista Electrónica de Investigación Educativa*, 2(1), 11-28. Recuperado de <https://redie.uabc.mx/redie>
- Brown, F. G. (1999). *Principios de la Medición en Psicología y Educación* (8ª ed.). México, DF: Manual Moderno.
- Brown, J. D. (1980). Relative merits of four methods for scoring cloze tests. *The Modern Language Journal*, 64(3), 311-317. doi: 10.1111/j.1540-4781.1980.tb05198.x
- Brown, J. D. (1988). Tailored cloze: Improved

- with classical item analysis techniques. *Language Testing*, 5(1), 19-31. doi: [10.1177/026553228800500102](https://doi.org/10.1177/026553228800500102)
- Brown, J. D. (1989). Cloze item difficulty. *Journal of the Japan Association of Language Teachers*, 11(1), 46-67. Recuperado de <http://jalt-publications.org/jj>
- Brown, J. D. (2013). My twenty-five years of cloze testing research: So what? *International Journal of Language Studies*, 7(1), 1-32. Recuperado de <http://www.ijls.net>
- Cairney, T. H. (2018). *Enseñanza de la comprensión lectora* (6ª ed.). Madrid, España: Morata.
- Calderón-Ibáñez, A., & Quijano-Peñuela, J. (2010). Características de comprensión lectora en estudiantes universitarios. *Revista Estudios Socio-Jurídicos*, 12(1), 337-364. Recuperado de <https://revistas.urosario.edu.co/index.php/sociojuridicos>
- Carlisle, J., & Rice, M. (2004). Assessment of reading comprehension. En A. Stone, E. Silliman, B. Ehren & K. Apel (Eds.), *Handbook of Language and Literacy: Development and disorders* (pp. 521-555). New York, NY: Guilford.
- Condemarin, M., & Milicic, N. (1988). *Test de cloze: Aplicaciones psicopedagógicas*. Santiago de Chile, Chile: Andrés Bello.
- Cortada de Kohan, N. (2004). Teoría de respuesta al ítem: Supuestos básicos. *Revista Evaluar*, 4(1), 95-110. Recuperado de <https://revistas.unc.edu.ar/index.php/revaluar>
- Costa, P. (2006). *Hábito de leitura e compreensão de textos: Uma análise da realidade de pós-graduados em Administração*. (Tesis de Maestría). Recuperado de repositorio.ufsm.br/handle/1/4735
- Cubo de Severino, L., & Bosio, I. V. (2011). La tesis como clase textual y su proceso de escritura. En L. Cubo de Severino, H. Puiatti, & N. Lacon (Eds.), *Escribir una Tesis. Manual de Estrategias de Producción* (pp. 13-36). Córdoba, Argentina: Comunicarte.
- Chino-Vilca, B., & Zegarra-Valdivia, J. (2019). Implicancia de las funciones ejecutivas “frías y calientes” en la comprensión lectora, en una muestra de estudiantes universitarios peruanos. *Psicología, Conocimiento y Sociedad*, 9(1), 59-82. Recuperado de <https://revista.psico.edu.uy/index.php/revpsicologia>
- Datta, S., & Macdonald-Ross, M. (2002). Reading skills and reading habits: A study of new Open University undergraduate reserves. *Open Learning*, 17(1), 69-88. doi: [10.1080/02680510120110193](https://doi.org/10.1080/02680510120110193)
- De Oliveira, K. L. (2011). Considerações acerca da compreensão em leitura no ensino superior. *Psicologia: Ciência e Profissão*, 31(4), 690-701. doi: [10.1590/S1414-98932011000400003](https://doi.org/10.1590/S1414-98932011000400003)
- Difabio de Anglat, H. (2008). El test cloze en la evaluación de la comprensión del texto informativo de nivel universitario. *Revista de Lingüística Teórica y Aplicada*, 46(1), 121-137. Recuperado de <http://www2.udec.cl/rla/principal.htm>
- Esquivel-Gámez, I., Martínez-Olvera, W., Córdoba del Valle, R., & Reyes-Gutiérrez, C. (2016). Memoria operativa y lectura comprensiva: Medición con pruebas de amplitud lectora y tipo cloze en ámbitos pre y universitarios. *Apertura. Revista de Innovación Educativa*, 8(2), 38-53. Recuperado

de <http://www.udgvirtual.udg.mx/apertura/index.php/apertura/index>

- Gellert, A., & Elbro, C. (2013). Cloze tests may be quick, but are they dirty? Development and preliminary validation of a cloze test of reading comprehension. *Journal of Psychoeducational Assessment, 31*(1), 16-28. doi: [10.1177/0734282912451971](https://doi.org/10.1177/0734282912451971)
- González-Moreyra, R. (1998). Comprensión lectora en estudiantes universitarios iniciales. *Persona, 1*, 43-65. Recuperado de <https://dialnet.unirioja.es/servlet/revista?codigo=12630>
- Green, B. (2001). Testing reading comprehension of theoretical discourse with cloze. *Journal of Research in Reading, 24*(1), 82-98. doi: [10.1111/1467-9817.00134](https://doi.org/10.1111/1467-9817.00134)
- Hernández-Sampieri, R., Fernández-Collado, C., & Baptista-Lucio, P. (2014). *Metodología de la Investigación* (6ª ed.). México, DF: McGraw-Hill.
- IBM Corporation. (2017). IBM SPSS Statistics for Windows (Versión 25.0) [Software de cómputo]. Armonk, NY: IBM.
- Jonz, J. (1990). Another turn in the conversation: What does cloze measure? *TESOL Quarterly, 24*(1), 61-83. doi: [10.2307/3586852](https://doi.org/10.2307/3586852)
- Kleijn, S. (2018). *Clozing in on readability: How linguistic features affect and predict text comprehension and on-line processing*. Utrecht, Países Bajos: LOT. Recuperado de <https://dspace.library.uu.nl/handle/1874/363346>
- Kleijn, S., Pander-Maat, H., & Sanders, T. (2019). Cloze testing for comprehension assessment: The HyTeC-cloze. *Language Testing, 36*(4), 553-572. doi: [10.1177/0265532219840382](https://doi.org/10.1177/0265532219840382)
- Kline, P. (1998). *The new psychometrics: Science, psychology and measurement*. Londres, Reino Unido: Routledge.
- Kobayashi, M. (2002). Cloze tests revisited: Exploring item characteristics with special attention to scoring methods. *The Modern Language Journal, 86*(4), 571-586. doi: [10.1111/1540-4781.00162](https://doi.org/10.1111/1540-4781.00162)
- León-Cascón, J. A. (2004). ¿Por qué las personas no comprenden lo que leen? *Psicología Educativa, 10*(2), 101-116. Recuperado de <https://journals.copmadrid.org/psed>
- Lozano-Fernández, L. M., & de la Fuente-Solana, E. I. (2015). Diseño y validación de cuestionarios. En E. Pantoja-Vallejo (Coord.), *Manual básico para la realización de tesis, tesis y trabajos de investigación* (2ª ed., pp. 247-272). Madrid, España: EOS.
- Márquez, C., Ibáñez, P., & Pérez, C. (2011). Concordancia entre aplicaciones de textos informativos y científicos elaborados mediante la técnica Cloze en alumnos universitarios de carreras del área de la Salud. *Revista de Educación en Ciencias de la Salud, 8*(2), 133-138. Recuperado de <http://www2.udec.cl/ofem/recs>
- Oller, J. W., & Jonz, J. (Eds.). (1994). *Cloze and coherence*. Lewisburg, PA: Bucknell University Press.
- O'Toole, J. M., & King, R. A. R. (2010). A matter of significance: Can sampling error invalidate cloze estimates of text readability? *Language Assessment Quarterly, 7*(4), 303-316. doi: [10.1080/15434303.2010.515699](https://doi.org/10.1080/15434303.2010.515699)
- O'Toole, J. M., & King, R. A. R. (2011). The deceptive mean: Conceptual scoring of cloze entries differentially advantages more able

- readers. *Language Testing*, 28(1), 127-144. doi: [10.1177/0265532210375687](https://doi.org/10.1177/0265532210375687)
- Parodi, G. (2001). Comprensión y producción lingüística: Una nueva mirada al procesamiento del discurso escrito. *Versión. Estudios de Comunicación y Política*, 11, 59-97. Recuperado de <http://version.xoc.uam.mx/index.php>
- Peña-Calvo, A. (1997). El Cloze como técnica para enseñar estrategias de comprensión lectora en una lengua extranjera (inglés). En Departamento de Filología Inglesa, Universidad de Murcia (Ed.), *Estudios de lingüística aplicada y literatura: Homenaje póstumo al profesor Juan Conesa Sánchez* (pp. 203-215). Murcia, España: Universidad de Murcia, Servicio de Publicaciones.
- Real Academia Española. (2010). *Nueva gramática de la lengua española*. Manual. Madrid, España: Espasa.
- Real Academia Española. (2019). *Corpus de Referencia del Español Actual -CREA-* [Base de datos]. Recuperado de <http://corpus.rae.es>
- Sadeghi, K. (2008). Measuring reading comprehension: The judgmental validity of cloze procedure. *Iranian Journal of Applied Linguistics*, 11(2), 115-132. Recuperado de https://ijal.khu.ac.ir/index.php?slc_lang=en&sid=1
- Shahnazari-Dorcheh, M., Roshan, S., & Hesabi, A. (2012). What is the optimum length of a cloze test? *International Journal of English Linguistics*, 2(5), 142-153. doi: [10.5539/ijel.v2n5p142](https://doi.org/10.5539/ijel.v2n5p142)
- Tabullo, A. J., Shalom, D., Sevilla, Y., Gattei, C. A., París, L., & Wainselboim, A. (2020). Reading comprehension and predictability effects on sentence processing: An event-related potential study. *Mind, Brain, and Education*, 14(1), 32-50. doi: [10.1111/mbe.12205](https://doi.org/10.1111/mbe.12205)
- Taylor, W. L. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism and Mass Communication Quarterly*, 30(4), 415-433. doi: [10.1177/107769905303000401](https://doi.org/10.1177/107769905303000401)
- Trace, J. (2020). Clozing the gap: How far do cloze items measure? *Language Testing*, 37(2), 235-253. doi: [10.1177/0265532219888617](https://doi.org/10.1177/0265532219888617)
- Trace, J., Brown, J. D., Janssen, G., & Kozhevnikova, L. (2017). Determining cloze item difficulty from item and passage characteristics across different learner backgrounds. *Language Testing*, 34(2), 151-174. doi: [10.1177/0265532215623581](https://doi.org/10.1177/0265532215623581)
- Watanabe, Y., & Koyama, D. (2008). A meta-analysis of second language cloze testing research. *Second Language Studies*, 26(2), 103-133. Recuperado de <https://scholarspace.manoa.hawaii.edu/handle/10125/40694?mode=full>

Anexo 1. Texto cloze

En el siguiente texto se han suprimido varias palabras para que las anticipes.

Debes colocar una palabra en cada blanco (destacándola con color); los términos omitidos, cualquiera sea su longitud, han sido reemplazados por una línea de extensión constante.

Nota: agregamos el número de cada blanco para facilitar la lectura de la clave de corrección; no se incluyen en el texto presentado a los participantes del taller.

Principios para un final: reflexiones en torno a la escritura de la tesis

La escritura, la presentación y la defensa de la tesis son las etapas conclusivas de los estudios de posgrado, que exigen al estudiante el ordenamiento del conocimiento que ha acumulado, analizado y desarrollado durante su carrera. Este 1. ----- lo compromete con dos 2. ----- interdependientes (de investigación y de 3. -----), de los cuales 4. ----- un producto final llamado 5. ----- . La escritura, por 6. ----- parte, junto con el 7. ----- del tema de investigación, 8. ----- el conocimiento de las 9. ----- discursivas necesarias para la 10. ----- de textos, en particular 11. ----- que permiten elaborar 12. ----- académicos (e.g. el proyecto de 13. -----, la monografía, la tesis). 14. -----, en el nivel de 15. -----, el estudiante, en su 16. ----- de escritor, debe resolver 17. ----- relativos tanto al contenido 18. ----- a la comunicación de 19. ----- contenido.

La escritura suele 20. ----- como una actividad que 21. ----- de un problema comunicativo 22. ----- resolución depende de la 23. ----- de un texto escrito. En este 24. -----, a fin de alcanzar con éxito la 25. ----- del problema, es necesario 26. ----- cada uno de los procesos 27. ----- en toda tarea de composición 28. -----: la planificación, la puesta en 29. ----- y la revisión, que 30. ----- alternan y se influyen 31. ----- durante todo el 32. -----.

La planificación 33. ----- una evaluación inicial de la situación 34. -----: tema, objetivos del 35. -----, destinatarios, formato o género, extensión, 36. ----- del que se dispone 37. ----- elaborarlo. Conduce a 38. ----- un plan de escritura, 39. ----- un índice comentado o una 40. ----- de tareas por hacer. 41. ----- evidente que detenerse en esta 42. ----- facilita la producción del texto. 43. ----- en discurso significa 44. ----- en palabras las ideas, 45. ----- como guía el plan de 46. ----- (más o menos 47. -----). Para llevar a 48. ----- esta tarea, es necesario 49. ----- las determinaciones impuestas 50. ----- la gramática de la lengua y el tipo de 51. ----- (e.g. vocabulario, sintaxis, frases frecuentes) 52. ----- por la situación comunicativa.

La 53. ----- comprende dos subprocesos: 54. ----- y edición.

En la evaluación, se 55. ----- la adecuación del texto 56. ----- a diferentes niveles: desde 57. ----- funcionamiento global 58. ----- la ortografía. Así, se 59. ----- que el texto se 60. ----- al problema planteado por la 61. ----- de escritura y al plan trazado, y su 62. ----- a las reglas gramaticales y 63. -----.

La edición busca 64. ----- los problemas del texto a 65. ----- local o global. Las operaciones de 66. ----- podrían describirse como 67. ----- de borrado o supresión de 68. -----, agregado o 69. -----, desplazamiento a lo largo del sintagma, etc.

Reflexionar 70. ----- estos procesos, 71. ----- conscientes nuestros propios 72. ----- mentales, es un punto de 73. ----- confiable para aprender a resolver los 74. ----- planteados por la escritura. A 75. ----- reflexión sobre la cognición, 76. -----, puede sumarse una 77. ----- sobre los usos 78. ----- del lenguaje y las 79. ----- epistemológicas de esos usos, es decir, la reflexión metalingüística. Ambas son instrumentos útiles para mejorar nuestra tarea como escritores académico-científicos.

Anexo 2. Clave de corrección

En el siguiente listado, la primera palabra (en letra itálica) corresponde al ítem *verbatim*; el resto, a las alternativas correctas ordenadas por frecuencia en la muestra; las casillas en color gris, a los ítems no discriminativos.

1. <i>ordenamiento</i> , proceso, trabajo, itinerario, quehacer
2. <i>procesos</i> , tareas, etapas/fases, actividades/acciones, objetivos/propósitos, instancias
3. <i>escritura</i> , redacción, comunicación
4. <i>resulta/resultará</i> , deviene
5. <i>tesis</i>
6. <i>su</i> , otra
7. <i>dominio</i> , desarrollo, estudio, conocimiento
8. <i>exige</i> , requiere, supone, necesita, implica, comprende, demanda, involucra
9. <i>estrategias</i> , herramientas, técnicas, formas, reglas, normas, estructuras, modalidades, prácticas
10. <i>producción</i> , redacción, elaboración, escritura, construcción, composición, estructuración
11. <i>aquellas</i> , las
12. <i>textos</i> , escritos, géneros
13. <i>investigación</i> , tesis
14. <i>Así</i> , Entonces, Ciertamente, Especialmente
15. <i>posgrado</i> , doctorado
16. <i>rol</i> , papel, carácter, función, condición
17. <i>aspectos</i> , problemas, asuntos, temas, tópicos, interrogantes
18. <i>como</i>
19. <i>dicho</i> , ese, este, tal
20. <i>concebirse</i> , presentarse, considerarse, entenderse, verse, definirse, aparecer, plantearse, pensarse, interpretarse, caracterizarse, proponerse, identificarse, comprenderse
21. <i>parte</i> , surge, resulta, procede, deriva, deviene
22. <i>cuya</i>
23. <i>elaboración</i> , producción, redacción, composición, construcción, formulación, realización, preparación
24. <i>sentido</i> , proceso, trayecto, camino, trabajo
25. <i>resolución</i> , solución
26. <i>realizar</i> , considerar, conocer, respetar, concretar, desarrollar, desplegar, entender, dominar, transitar, recorrer, atravesar, aplicar

27. <i>involucrados</i> , implicados, presentes, contenidos
28. <i>escrita</i> , textual
29. <i>discurso</i> , palabras, texto
30. <i>se</i>
31. <i>mutuamente</i> , dinámicamente, recíprocamente, respectivamente
32. <i>proceso</i>
33. <i>implica</i> , es, requiere, supone, constituye, conlleva, involucra, representa, proporciona, contiene, comprende, prevé
34. <i>comunicativa</i> , comunicacional, escritural
35. <i>texto</i> , trabajo, escrito, autor, investigador, discurso, escritor, comunicador, emisor
36. <i>tiempo</i> , material
37. <i>para</i>
38. <i>producir</i> , realizar, diseñar, establecer, desarrollar, generar, trazar, armar, formular, esbozar, pensar, confeccionar, concretar, definir, estructurar, avanzar
39. <i>como</i> , con, elaborando
40. <i>lista</i> , enumeración, guía, grilla, agenda, tabla, sistematización, descripción
41. <i>Es</i> , resulta, parece
42. <i>etapa</i> , tarea, instancia, actividad, fase
43. <i>Puesta/poner</i>
44. <i>expresar</i> , poner, plasmar, transformar, traducir, trasponer, presentar, comunicar, materializar
45. <i>considerando</i> , teniendo, siguiendo, utilizando
46. <i>escritura</i> , texto
47. <i>detallado</i> , definido, preciso, explícito/explicitado, elaborado, delimitado, desarrollado, planificado, programado, estructurado, especificado, esbozado, bosquejado, completo
48. <i>cabo</i>
49. <i>conocer</i> , respetar, considerar, seguir, cumplir, revisar, comprender
50. <i>por</i> , desde
51. <i>lenguaje</i> , discurso, registro, recursos, género, código
52. <i>requerido</i> , exigido, determinado, impuesto, establecido, definido, demandado, delimitado, condicionado
53. <i>revisión</i>
54. <i>evaluación</i>
55. <i>verifica</i> , analiza, observa, revisa, valora, examina, comprueba, considera, controla, determina, pondera, juzga

56. <i>elaborado</i> , escrito, producido, realizado, redactado, construido, final
57. <i>su</i>
58. <i>hasta</i> , a
59. <i>corroborar</i> , evalúa, asegura, comprueba, verifica, busca, controla, garantiza, confirma, establece, valora
60. <i>ajuste</i> , adecue, adapte, atenga, convenga
61. <i>tarea</i> , situación, actividad
62. <i>adecuación</i> , ajuste, sujeción, fidelidad
63. <i>ortográficas</i> , formales, lingüísticas
64. <i>solucionar</i> , resolver, corregir, subsanar, superar, acomodar, eliminar, despejar
65. <i>nivel</i> , escala
66. <i>edición</i>
67. <i>operaciones</i> , tareas, acción/es, proceso/s, actividades, las, procedimientos, técnicas, trabajo, ejercicio, labor, estrategias, rutinas, aplicación, aquellas
68. <i>elementos</i> , palabras, texto, errores, frases, contenido, información, términos, erratas, fragmentos, ideas
69. <i>inserción</i> , modificación/es, adición, reemplazo, sustitución, ampliación, cambio/s, aclaración, incorporación, mejora/s, corrección, expansión, reducción, explicitación, ajustes, inclusión
70. <i>sobre</i> , en
71. <i>hacer</i> , haciendo, volver, tornar, volviendo
72. <i>procesos</i> , esquemas, mecanismos, procedimientos, estados
73. <i>partida</i>
74. <i>problemas</i> , desafíos, retos, inconvenientes, obstáculos
75. <i>esta</i> , nuestra, esa
76. <i>además</i> , también, asimismo, finalmente, conjuntamente, ciertamente
77. <i>reflexión</i> , consideración, mirada, más, deliberación, discusión, concientización, evaluación, valoración, exploración, similar
78. <i>académicos</i> , científicos, especializados
79. <i>razones</i> , bases, consecuencias, reglas, implicancias, normas, raíces, derivaciones, fundamentaciones, condiciones, especificaciones, fuentes, pautas, categorías