

Evaluación de la comprensión de lectura en lengua extranjera

María Aurora Marrón Orozco

Introducción

¿Qué medimos cuando decimos que evaluamos la comprensión de un texto en lengua extranjera (L2)? ¿Qué problemas enfrenta el alumno cuando lee en L2, problemas de lectura o problemas de conocimientos de L2? ¿Qué debe medir un examen, habilidades de lectura o conocimientos de L2?

El consenso de las investigaciones es bastante claro. Los lectores en L2 necesitan pasar un *nivel umbral* de conocimientos de L2 antes de que sus habilidades de lector en su lengua materna puedan transferirse a la lectura en L2. Este nivel umbral interactúa con conocimientos previos y con el texto de tal manera que con determinados textos de ciertos temas, el lector necesita menos dominio de la L2 que con otros textos y temas. Una implicación obvia de las investigaciones es que el lector principiante en L2 necesita mejorar y aumentar sus conocimientos de L2 para poder leer bien (Alderson, 2000).

Con esta problemática en mente, en este capítulo se incluyen principios y lineamientos que el profesor deberá tomar en cuenta cuando se proponga evaluar a sus alumnos de un curso de comprensión de lectura (CL) en L2, las diferentes maneras y los diversos propósitos para evaluar durante un curso, las etapas en el proceso de diseñar una prueba o examen y algunas recomendaciones prácticas para la etapa de elaboración de reactivos de determinados instrumentos de evaluación.

Principios y lineamientos para evaluar CL en L2

Varios investigadores como Nuttall (1996) y Grabe (1991) han propuesto principios y lineamientos que pueden ser de mucha utilidad para los profesores cuando necesitan diseñar una prueba o un examen de comprensión de lectura en L2 para sus alumnos y tienen las dudas que acechan a cual-

quier diseñador de exámenes: ¿Cómo ser justo? ¿Cómo debe ser mi examen para que sea válido?

Bachman y Palmer (1996) recomiendan a los profesores que tengan en mente las siguientes seis consideraciones filosóficas:

- Relacionar la evaluación de comprensión de lectura con la enseñanza-aprendizaje de comprensión de lectura y con el uso real que el alumno hace de sus lecturas.
- Diseñar sus exámenes para estimular que la capacidad de leer de su alumno que toma el examen se note y sea evidente para él y para el profesor.
- Considerar lo que es ‘justo’ para su alumno.
- Humanizar el proceso de evaluar. Es decir, encontrar maneras de involucrar a sus alumnos de manera directa en el proceso de evaluar; tratar a los examinados como individuos responsables; darles la mayor cantidad de información sobre el proceso de evaluación.
- Pedir responsabilidad a los demás y a sí mismo sobre el uso que hagan de los exámenes.
- Reconocer que las decisiones tomadas con base en resultados de exámenes están llenas de dilemas y que no hay respuestas universales a estos dilemas.

Alderson (2000) señala los siguientes lineamientos para evaluar comprensión de lectura en L2:

- La comprensión de lectura debe ser examinada dentro de una batería de textos enfocados en contenido; textos que tengan sentido para los lectores, textos que les interesen, que se relacionen con sus antecedentes académicos, con sus intereses en su tiempo libre, que sean del nivel adecuado para los lectores, etc., es decir, textos que motiven una lectura a profundidad. No es conveniente usar textos anodinos, artificiales y sin ningún interés para el alumno.
- Los estudiantes deben ser evaluados en un rango de estrategias y habilidades relevantes y sus resultados pueden darse en un formato que retroalimente al alumno y diagnostique sus fortalezas y debilidades.
- Las tareas solicitadas en el examen deben poder realizarse en el tiempo disponible para no frustrar a los alumnos.

- Debe reconocerse que los conocimientos previos influyen en toda comprensión y, por lo tanto, es bueno todo intento para permitir que los conocimientos previos faciliten la ejecución en el examen.
- Los exámenes de comprensión de lectura pueden estar abiertos a la posibilidad de interpretaciones múltiples. Los diseñadores deben ser flexibles y aceptar un rango posible de diferentes interpretaciones.
- Para evaluaciones continuas se pueden formar grupos de discusión sobre posibles interpretaciones de un texto. Los profesores pueden mantener un registro de los diversos puntos que salieron en la discusión.
- La lectura en bulto o extensiva no debe dejarse a un lado en los procedimientos de evaluación. Se pueden evaluar los portafolios de los alumnos con los textos leídos y las tareas realizadas con estos, además de las opiniones del lector sobre cada lectura. El profesor puede llevar un registro de cada lectura que hacen sus alumnos.
- La velocidad de lectura no debe evaluarse a menos que se evalúe también la comprensión, pero sí debe tomarse en cuenta porque muchas veces evaluamos comprensión y no le damos importancia al tiempo que el alumno tardó en comprender.
- Se necesita más investigación para encontrar maneras de evaluar habilidades específicas como la habilidad de síntesis o los conocimientos metacognitivos. A veces consideramos suficiente evaluar la lectura de una manera global con la esperanza de que el alumno haya usado estas habilidades al leer. Sin embargo, más a menudo los diseñadores de exámenes se enfocan a la evaluación de: a) vocabulario, b) estructuras sintácticas, c) habilidades discursivas y d) habilidades para inferir el significado de palabras. Eso está bien pero hay otro tipo de habilidades, que pueden considerarse de orden superior, como por ejemplo identificar la ironía del autor, que los diseñadores de exámenes muchas veces olvidan evaluar.
- Es muy importante que los diseñadores de exámenes consideren hasta qué punto sus exámenes reflejan el proceso de aprendizaje de lectura en L2 y no solo el producto de la comprensión de un texto dado.

Propósitos para evaluar CL en L2

Enumeramos una serie de reflexiones filosóficas sobre el acto de evaluar y una serie de lineamientos y observaciones alrededor de la evaluación de la comprensión de lectura en L2. Ahora pensemos en los diferentes propó-

sitos que tenemos para evaluar a nuestros estudiantes y en las diversas maneras en que podemos hacerlo. Díaz Barriga y Hernández Rojas (2002) dividen las actividades o tareas de evaluación que se incluyen en un curso y de las cuales el profesor puede llevar un registro, en:

- *Informales*, como la participación en clase o contestar preguntas orales y escritas durante la clase. El profesor puede llevar un diario de clase y anotar sus observaciones sobre el desempeño de sus alumnos en cada una (véase el Capítulo 14).
- *Semiformales*, como un conjunto de reportes, trabajos y ejercicios de cada lectura hecha en el salón de clases o de tareas en casa a lo largo del curso o un *portafolio* que cada alumno elaborará y organizará durante el curso. En estas tareas de evaluación semiformales se puede fomentar la evaluación entre compañeros de grupo y la autoevaluación (véase el Capítulo 14).
- *Formales*, como pruebas o exámenes o desempeñar alguna tarea de lectura auténtica que se pueda evaluar con una escala o con una rúbrica. Las *rúbricas*, según Díaz Barriga y Hernández Rojas (2002), son guías de puntaje que permiten describir el grado en el cual un aprendiente está ejecutando un proceso o un producto. La rúbrica o método para calificar debe establecerse al diseñar el instrumento de evaluación y el profesor debe informar a quienes harán el examen sobre la forma de calificar y lo que se tomará en cuenta. Bachman y Palmer (1996) ofrecen los siguientes ejemplos de tareas que se pueden calificar con rúbricas establecidas previamente y que están relacionadas con el uso real de la lectura en L2: visitar un lugar y leer un folleto informativo sobre el mismo y planear un itinerario de visitas, o consultar un horario de TV para localizar programas de determinado tema o interés.

Las razones o propósitos que el profesor tiene para llevar a cabo evaluaciones informales, semiformales y formales durante su curso pueden ser de tres tipos:

- *Diagnóstico*: es decir, para saber si el alumno tiene conocimiento de algún o algunos puntos y/o temas o hizo uso de estrategias específicas, y así poder iniciar o continuar su formación a partir de lo que ya sabe.

- *Formativos*: para ayudar al alumno a autoevaluarse, para determinar si ya entendió o ya puede aplicar lo que el profesor enseñó o lo que contiene el programa del curso. También para saber si necesita ayuda y más trabajo y para dar seguimiento al progreso de los estudiantes durante el proceso de aprendizaje.
- *Sumativos*: para otorgar una calificación al alumno y, en su caso, promoverlo o darle una certificación y también para determinar la eficiencia del programa.

Elaborando un poco más en lo anterior, dentro de un enfoque comunicativo –y más recientemente en el aprendizaje basado en tareas–, la condición principal es que las tareas incluidas en un instrumento de evaluación reflejen tanto como sea posible el procesamiento real del discurso y cubran el rango de habilidades que contribuyen a leer bien. Como se estableció en el párrafo anterior, cuando diseñemos estas tareas debemos pensar también en el propósito que tenemos para evaluar a nuestros alumnos. Por ejemplo, diseñaremos una prueba que nos ayudará a ‘diagnosticar’ la situación en comprensión de lectura de nuestros alumnos y que tendrá un propósito de evaluación ‘formativo’, si queremos saber si están poniendo en práctica algunas estrategias que hemos estudiado en clase o queremos saber si pueden identificar las ideas principales en un texto, o si pueden escribir un resumen del texto. Esta prueba tendrá el objetivo de diagnosticar la situación de cada uno de ellos con respecto a estos temas del programa y proponerles tareas y ejercicios alternativos extras para ayudarles a mejorar su lectura y su desempeño en el curso. Por otro lado, si queremos saber si nuestros alumnos dominan los mismos temas del programa mencionados anteriormente pero nuestro objetivo es otorgar una calificación al final de una unidad, al final de un periodo o al final del curso, entonces se trata de una evaluación ‘sanativa’ y el diseño de las tareas deberá ser más cuidadoso y riguroso porque los resultados de esta prueba o examen son, generalmente, determinantes para los alumnos.

Maneras de evaluar CL en L2

Detengámonos ahora en el proceso mismo de diseñar una prueba o examen de comprensión de lectura en L2 para nuestros alumnos. Los pasos en la planeación de un examen son los siguientes:

- Determinar el propósito del examen.
- Especificar los objetivos del examen.
- Definir el contenido.
- Determinar los pesos (o valores) relativos de las diferentes partes.
- Determinar el (los) método(s) y procedimientos a utilizarse; congruentes con el objetivo y el contenido.

Alderson, Clapham y Wall (1995) proponen concentrar toda la información de estos cinco pasos en una *tabla de especificaciones*, que es una especie de anteproyecto del examen que precisa lo que el examen evalúa y cómo lo hace. Estas especificaciones son necesarias para quienes diseñan el examen, para quienes lo validan y para los usuarios del examen. La tabla de especificaciones para un examen de comprensión de lectura en L2 deberá incluir entonces:

- El propósito del examen: ¿qué tipo de examen es: aprovechamiento, diagnóstico, certificación?
- Una descripción de la población que tomará el examen (la población meta). Sus características: edad, nivel de conocimientos lingüísticos, razones para tomar el examen e intereses profesionales.
- Tipos de textos que se utilizarán: temas, fuentes, grado de autenticidad, longitud, dificultad, funciones que representan, complejidad del lenguaje.
- Las habilidades y estrategias a evaluar. Por ejemplo, buscar la idea principal, buscar detalles específicos, inferir significados del contexto, etc.
- Elementos lingüísticos que se evaluarán: vocabulario, cohesión, etc.
- ¿Qué tipo de tareas se le pedirá al alumno que haga?
- ¿Cuántas secciones contendrá el examen? ¿Qué tiempo tomará resolver cada una de ellas? ¿Cuánto valdrá cada sección? ¿Cuánto valdrá el examen completo?
- ¿Cuántas preguntas o reactivos se incluyen en cada sección? ¿Qué peso o valor se le da a cada uno de ellos?
- ¿De qué tipo serán las preguntas o reactivos? Respuesta corta, relacionar columnas, *cloze*, resumen, mapa conceptual, mapa mental, verdadero/falso, opción múltiple, etc.
- Instrucciones para los alumnos: ¿en lengua materna o extranjera?, ¿con ejemplos?

Aunque estos factores aparecen de manera lineal, la tabla de especificaciones puede desarrollarse en forma cíclica y se puede ir haciendo a medida que se va planeando y diseñando el examen.

Una vez que tenemos claro nuestro propósito para evaluar y las características de la población que tomará el examen porque son nuestros alumnos, lo primero que debemos hacer es seleccionar el texto o textos adecuados, respetando los criterios mencionados en el primer lineamiento de la tabla y recordando los criterios de *exploitability*. A continuación debemos pensar en las tareas, actividades o preguntas que les pediremos a nuestros alumnos que hagan con ese texto o textos para que nos demuestren que han comprendido su mensaje y contenido.

Las tareas, actividades o preguntas que podemos incluir en un examen de comprensión de lectura en L2 pueden ser de muchos tipos y deben ser similares a las actividades que han realizado nuestros alumnos durante el curso. A continuación se presenta la Tabla 1 con algunas de las técnicas más usadas y aspectos como ‘dificultad de elaboración’, ‘realismo’ y ‘ventajas’, que debemos considerar cuando las usemos.

Tabla 1. Tipos de exámenes: ventajas y dificultades

| | Opción múltiple | Pregunta con respuesta corta | Transferencia de información | Verdadero/falso |
|---------------------------|---|--|---|---|
| Dificultad de elaboración | Muy difícil. Se requiere entrenamiento para escribirlas. | Fácil de escribir. Cuidar formulación. | Difícil. Llenar diagramas, cuadros, secuencia de eventos, mapas conceptuales. | Difícil. Para evaluación sumativa se deben elaborar muchos reactivos. |
| Realismo de la pregunta | Poco. La vida real no presenta tantas alternativas. | Mucho. Al leer los estudiantes se hacen preguntas continuamente. | Mucho. Los estudiantes estudian procesos, diagramas y mapas conceptuales. | Algo. En la vida real decidimos si algo es verdadero o falso. |
| Ventaja más importante | La objetividad para calificarla. Es confiable, simple y rápida para calificar. Ideal para evaluaciones sumativas de grupos grandes. | El lector tiene que encontrar la respuesta y escribirla, en L1 o L2. | Lo auténtico de la tarea en un medio académico. | La objetividad para calificarla. Es rápida para calificar. |

Cuando hablamos de examen de opción múltiple nos referimos a los componentes que lo constituyen de la siguiente manera. Los exámenes de opción múltiple están formados por reactivos. Un reactivo es una unidad de medida que consiste en una pregunta o instrucción que requiere una respuesta del examinado. Cada reactivo está compuesto por: 1) un enunciado, que es el planteamiento que da lugar a una respuesta, 2) la respuesta correcta y 3) varios distractores, que son las opciones que representan respuestas incorrectas. Es decir, un enunciado presenta un problema o pregunta y puede elaborarse en forma de afirmación declarativa o interrogativa, seguido de tres a cinco opciones de respuesta, de las cuales una es la correcta. Las opciones de respuesta deben consistir en oraciones, palabras aisladas o soluciones a problemas de índole diversa.

Elaborar exámenes de opción múltiple es muy difícil y requieren un proceso de validación arduo antes de aplicarse a la población para la que fueron diseñados. Por estas razones, y porque contestar este tipo de opciones no es una actividad de lectura real, no se recomienda a los profesores elaborar este tipo de exámenes para sus grupos de comprensión de lectura. Otras técnicas y maneras de evaluar a sus alumnos se asemejan más a actividades de lectura en el mundo real y les darán más información de su aprovechamiento y avances en el curso tales como: preguntas abiertas, llenado de mapas conceptuales o diagramas, elaboración de resúmenes, verdadero/falso, relacionar columnas, *cloze*, etc.

Al igual que en la técnica de opción múltiple, a las preguntas que forman un examen de verdadero/falso se les llama reactivos. Estos reactivos están formados por un enunciado y la opción de marcar verdadero o falso. El reactivo de verdadero/falso (V/F) es muy solicitado, porque se piensa que es fácil de elaborar pero en realidad no es así. Este tipo de reactivo se presta especialmente para estimar un conocimiento de hechos que son, inequívocamente, verdaderos o falsos. Es idóneo para situaciones en las que existen únicamente dos opciones opuestas entre sí. Cuando redactamos reactivos de V/F debemos asegurarnos que el reactivo pueda clasificarse como verdadero o falso. No debemos emplear términos ambiguos o indefinidos para designar grado o cantidad como ‘frecuentemente’, ‘a veces’, ‘en la mayoría de los casos’ porque tienen sentidos diferentes para los lectores. Es mejor evitar palabras en el enunciado, como: ‘todos’, ‘nunca’ y ‘ningún’ ya que dan lugar a generalizaciones tan amplias que suelen ser falsas. Tampoco debemos utilizar enunciados negativos o negaciones

dobles. El enunciado negativo requiere un proceso de razonamiento un tanto enredado para descifrar su significado. Como el alumno tiene 50% de oportunidad de adivinar, se recomienda que las pruebas de V/F tengan un número de preguntas relativamente grande cuando tienen un propósito sumativo.

Hay otras técnicas que se usan para evaluar la comprensión de lectura en lengua materna y que también se han usado para L2, como el *cloze* en diversas modalidades. Para elaborar un examen *cloze* a partir de un texto dado se omiten palabras de manera mecánica a partir del segundo o tercer párrafo del texto. Puede omitirse una palabra de cada cinco palabras, cada seis palabras, cada siete... hasta cada 11 palabras. Es importante dejar las primeras oraciones o el primero y segundo párrafos sin huecos y luego decidir cada cuántas palabras se omitirá una en el resto del texto. El *cloze* es una prueba integrativa, es decir, es una prueba que evalúa conocimientos lingüísticos de la L2: gramaticales, sintácticos y discursivos y habilidades y estrategias para leer. Pero es complicado determinar qué evalúa cada reactivo o conjunto de reactivos. Un *cloze* diseñado como se indicó arriba es un examen fácil de elaborar pero muy difícil de calificar, ya que el profesor tendrá que decidir si acepta como respuesta correcta solo la palabra que el autor del texto usó en cada hueco o acepta sinónimos de las palabras omitidas; si acepta palabras en L1 o solo en L2, etc. En grupos pequeños esto se puede hacer y permitir evaluación formativa entre compañeros y entre profesor y alumnos. Cuando el examen *cloze* se usa para poblaciones grandes con fines sumativos, a menudo se combina con reactivos de opción múltiple, en donde cada reactivo contiene la palabra correcta de cada hueco y de dos a cuatro palabras que funcionan como distractores.

Urquhart y Weir (1998) favorecen y recomiendan el uso del '*cloze* racional', que es una modalidad en la que el diseñador selecciona las palabras que se omitirán de acuerdo a lo que el profesor sabe que es difícil o fácil de la L2, la dificultad del texto, el curso en el que se aplicará, si es el caso, y el nivel de conocimientos que tenga la población que tomará el examen. En este tipo de prueba es menos difícil determinar qué es lo que cada reactivo evalúa.

Por último, las siguientes recomendaciones prácticas pueden ser muy útiles cuando diseñemos nuestros exámenes:

- El lenguaje que usemos para las preguntas, el vocabulario que contengan los reactivos y sus opciones son todos muy importantes. En ningún caso las preguntas o reactivos pueden contener vocabulario más difícil que el texto.
- Las preguntas o reactivos deben tener una sola respuesta correcta.
- Si el alumno entiende el texto debe poder contestar la pregunta.
- No se debe preguntar sobre información irrelevante del texto.
- Las preguntas no deben enfocarse a puntos gramaticales.
- El orden de las preguntas debe respetar el orden en que la información aparece en el texto.
- Las preguntas sobre actividades globales como actividades de predicción o *skimming* deben ir al principio. Nuttall (1996) sugiere actividades que llevan primero una mirada global del texto para después hacer una lectura detallada.
- No elaborar reactivos o preguntas que se puedan contestar sin leer el texto.
- El alumno tiene que leer el texto para contestar las preguntas.
- Pedir siempre a dos o más colegas que resuelvan el examen y nos hagan comentarios sobre el mismo antes de aplicarlo a nuestros alumnos.

Para concluir, es importante recordar que la selección del texto es uno de los aspectos más importantes tanto para propósitos de enseñanza como de evaluación, por lo que no debemos olvidar las características que debe poseer (véase el Capítulo 5).

Preguntas para reflexionar

- ¿Qué tipo de evaluación realizo?
- ¿Cuál es el propósito de la evaluación que realizo en mi clase?
- ¿Cómo podría mejorar los exámenes que elaboro?

Referencias

- Alderson, C. (2000). *Assessing reading*. Cambridge: Cambridge University Press.
- Alderson, J. C., Clapman, C. & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
- Bachman, L. F. & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.

- Cohen, A. (1994). *Assessing language ability in the classroom*. Boston: Heinle & Heinle.
- Díaz Barriga, F. & Hernández Rojas, G. (2002). *Estrategias docentes para un aprendizaje significativo*. (Cap. 8. Constructivismo y evaluación psicoeducativa). México: McGraw-Hill.
- Grabe, W. (1991). Current developments in second-language reading research. *TESOL Quarterly*, 25(3), 375-406.
- Nuttall, C. (1996). *Teaching reading skills in a foreign language*. Londres: Heinemann Educational.
- Urquhart, S. & Weir, C. (1998). *Reading in a second language. Process, product & practice*. *Applied Linguistics and Language Study*. Londres: Longman.