



LAS PRUEBAS DE CIERRE EN LA ENSEÑANZA DE LENGUAS EXTRANJERAS

THE CLOSING TESTS IN FOREIGN LANGUAGE TEACHING

M^a de los Milagros Esteban García*

RESUMEN

A pesar del tiempo transcurrido desde su creación, las pruebas de cierre se pueden considerar un instrumento válido, práctico y fiable en la enseñanza de lenguas extranjeras, y más concretamente en la evaluación del Inglés. Este artículo se remonta a los orígenes de la técnica, expone los distintos tipos de *cloze* y describe sus características más significativas. Finalmente destaca la vigencia de las pruebas de cierre y su valiosa aportación para el profesorado de idiomas, no sólo como instrumento de evaluación, sino también por su utilidad como actividad de aprendizaje en el aula.

Palabras clave: evaluación, pruebas de cierre, inglés como lengua extranjera.

ABSTRACT

In spite of the time gone by since its creation, closing tests can be considered a valid, practical and trustworthy instrument for foreign languages education, and more specifically in the evaluation of English. This article goes back to the origins of the technique, shows the different types of *cloze* and describes its most significant features. Finally it highlights the validity of closing tests and their valuable contribution to language teachers, not only as an evaluation tool, but also due to their usefulness as a learning activity within the classroom.

Key words: evaluation, closing tests, English as a foreign language.

* Doctora en Filología inglesa. Profesora Asociada en la Sección Departamental de Filología Inglesa de la Facultad de Ciencias Económicas y Empresariales. Universidad Complutense de Madrid.

INTRODUCCIÓN

Es más que probable que todos los que hemos abordado alguna vez el aprendizaje de una lengua extranjera nos hayamos enfrentado a una prueba de cierre en un momento u otro de ese proceso. Pero es también probable que ni siquiera nos hayamos preguntado el porqué de este tipo de prueba, o cómo mide nuestra competencia lingüística lo que, a primera vista, podría parecer un simple rompecabezas.

Partiendo del interés por encontrar instrumentos prácticos, válidos y fiables para la evaluación del inglés como Lengua Extranjera, este artículo se remonta a los orígenes de la técnica de cierre, creada por Taylor en los años 50, con clara influencia de las teorías de la Psicología de la Gestalt sobre la percepción de modelos visuales. Pretende además describir qué es una prueba de cierre, sus características y tipos, y destacar el importante papel que desempeñan actualmente como instrumento de evaluación en la enseñanza de Lenguas Extranjeras, especialmente del inglés.

A pesar de que los *clozes* nacieron hace ya más de medio siglo, su vigencia es indiscutible. La técnica de cierre, en cualquiera de sus formatos, ofrece muchas posibilidades al profesorado de idiomas. Aunque en este artículo nos centraremos en su utilización como instrumento de evaluación, también puede ser utilizado directamente en el aula como actividad de aprendizaje.

1. CONCEPTO DE «PRUEBA DE CIERRE» O CLOZE TECHNIQUE

La técnica «de cierre» o *cloze technique* fue creada por Wilson L. Taylor en 1953. Su aparición supuso el comienzo de un nuevo procedimiento de diseño de pruebas integradoras para la evaluación de la competencia lingüística.

Consiste en la mutilación de un texto mediante la omisión de un número determinado de sus elementos. La tarea que se propone al alumno es la recuperación del texto original:

A cloze unit may be defined as: any single occurrence of a successful attempt to reproduce accurately a part deleted from a «message» (any language product), by deciding from the context that remains, what the missing part should be (Taylor, 1953: 416).

El propio autor acuñó el nombre de *cloze* para estas pruebas. Según Oller (1979: 341), con esta denominación aludía directamente a *close*¹ («cerrar» en lengua inglesa), porque al completar los huecos de un texto que ha sido previamente mutilado se realiza un acto de *closure* o cierre, de modo semejante a lo que ocurre en la percepción de modelos visuales incompletos, según la Psicología de la Forma o Gestalt.

Por razones de espacio, no procede repasar minuciosamente en este artículo los principales rasgos de esta escuela psicológica² ni sus importantes aportaciones en el campo del aprendizaje. Tan sólo recordaremos algunos aspectos que resultan de especial interés para explicar las pruebas de cierre.

Para la Psicología de la Gestalt un todo «es una configuración compleja cuyos componentes se encuentran relacionados entre sí [...] todos los componentes adquieren significación dentro de la estructura global» (Moya Santoyo, 2002: 45). Uno de sus principios básicos es el de «pregnancia»; la tendencia a percibir los objetos como totalidades bien estructuradas y de la forma más simple. Este fenómeno se explica gracias a cuatro procesos internos: cierre, proximidad, continuidad y semejanza. El principio de cierre es el que inspiró a Taylor para la creación de los *clozes*. Según esta regla, las estructuras cerradas se perciben más fácilmente como unidades, pues tendemos a cerrar las configuraciones incompletas y a recordar como cerrado aquello que no lo está totalmente.

Este principio, que se observa claramente en la percepción visual, se puede aplicar también a otros campos. De manera semejante, según Taylor (1953), tendemos a completar un texto mutilado o incompleto, porque constituye un todo o unidad. Los elementos de un texto se interrelacionan y adquieren su pleno significado en la estructura global. La redundancia de la lengua y la gramática de expectativas nos ayudan a completar esa estructura, es decir, el todo.

La técnica de cierre resultó ser un procedimiento de evaluación prometedor y posteriormente ha sido muy popular en la enseñanza e investigación de Lenguas Extranjeras, pues numerosas investiga-

¹ Oller (1979: 341) se refiere al origen del término *cloze* y apunta lo siguiente: «The term is a mnemonic or perhaps a humorless pun intended to call to mind the process of closure celebrated by Gestalt psychologists». La palabra *cloze* es simplemente «a spelling corruption of the word *close*».

² La escuela psicológica gestáltica comenzó en Frankfurt am Main en 1910-12 con los estudios de Wertheimer.

ciones han puesto de manifiesto que es un instrumento práctico, válido y fiable, además de producir un efecto rebote positivo en la enseñanza.

2. FUNCIONAMIENTO DE LA TÉCNICA DE CIERRE

A continuación veremos qué mecanismos se ponen en funcionamiento para que el hablante de una lengua consiga «rellenar» los huecos que distorsionan su percepción global del texto, y cómo actúa el aprendiz de una segunda lengua, en nuestro caso, el Inglés.

La técnica de cierre parte del reconocimiento de la lengua como sistema funcional y creativo que contiene abundantes redundancias (Spolsky, 1973; Read, 1982 citado en Weir, 1990).

Spolsky (1973) enunció el principio de redundancia reducida de la lengua (*Reduced Redundance Principle* o RRP), que hace posible la comprensión de mensajes aunque algunos de sus elementos no aparezcan. En este principio se basan las pruebas de redundancia reducida, entre las cuales se incluyen las de cierre: «(they) test a subject's ability to function with a second language when noise is added or when portions of a test are masked» (Spolsky, 1973: 175).

Según Oller (1979: 344), el procedimiento de cierre mide la interiorización de los conocimientos gramaticales, ya que para predecir una palabra de un texto hay que utilizar las habilidades que subyacen a la actuación lingüística y así se demuestra el grado de competencia en la lengua. Cuando pedimos a un alumno que recupere una palabra omitida en un texto, hacemos que ponga en práctica esas habilidades.

Las posibilidades en cada punto del texto son limitadas. Para encontrarlas, el sujeto cuenta con cierta información; como las claves textuales y contextuales, ha de utilizar la redundancia de la lengua y el sistema de expectativas (*expectancy system*) que tiene como hablante. Para completar un texto correctamente hay que entender el texto y el contexto extralingüístico. Es necesario utilizar las claves y limitaciones sintácticas, morfológicas y semánticas que impone el sistema de la lengua, pero también inferir información del contexto extralingüístico. El buen conocimiento de la lengua supondrá una gran ayuda en este proceso.

The principle of reduced redundancy used in Information Theory is also thought to be involved, because the cloze test reduces natural

linguistic redundancies and requires the test taker to rely upon organizational constraints to fill in the blanks and infer meaning (Fotos, 1991: 315).

Taylor (1953) expresaba ya las ideas de gramática de expectativas y redundancia de la lengua, que constituyen la base de la técnica, y fueron posteriormente desarrolladas por Oller (1979) y Spolsky (1973) respectivamente:

Some words are more likely than others to appear in certain patterns or sequences. «Merry Christmas» is a more probable combination than «Merry birthday».

«Man coming» means the same as «A man is coming this way now». The latter, which is more like ordinary English, is redundant; it indicates the singular number of the subject three times (by «a», «man», «is») [...] Such repetitions of meaning, such internal ties between words, make it possible to replace «is», «this», «way», or «now», should any of them be missed (Taylor, 1953: 418).

Mostramos un ejemplo práctico del proceso que se sigue para completar las omisiones de un texto en lengua inglesa, teniendo en cuenta las limitaciones (*constraints*) que impone el propio sistema de la lengua (sintácticas, morfológicas, semánticas), y utilizando todo tipo de información lingüística y extralingüística.

Judith Taylor talks about her life as a top model.

I am sure that _____ (1) people think that the _____ (2) of a model is _____ (3) easy and very exciting. _____ (4) is true that I _____ (5) to some fantastic places _____ (6) I meet some interesting _____ (7). And the clothes, of _____ (8)... I love wearing beautiful _____ (9)!

Éste es el aspecto que presenta una prueba de cierre tradicional, de ratio fija. El alumno recibe un texto, adecuado a su nivel de competencia en la lengua meta (el inglés), con un determinado número de omisiones. Para recuperar el texto original debe rellenar esos huecos aprovechando las claves de que dispone.

En primer lugar, ha de tener en cuenta el tipo de texto (género, tema, contexto, etc.), puesto que cada género muestra determinadas convenciones y rasgos estilísticos. En esta ocasión, el texto es sencillamente el relato del estilo de vida de una modelo, contado por ella misma. Esta información inicial ya nos da una idea del registro de lengua, del tipo de vocabulario y estructuras gramaticales que probablemente aparecerán en el texto (términos relacionados con la vida

diaria, frente a otras posibilidades que encontraríamos en textos científicos, legales, literarios, etc.). Así comenzará a aplicar su gramática pragmática de expectativas.

Además, ha de valorar las limitaciones morfosintácticas y las claves que aporta la redundancia de la lengua. Se percatará de que probablemente la primera palabra omitida (1. *most*) sea un adjetivo o un determinante debido a su posición en la oración, delante de un nombre, mientras que la segunda (2. *lifestyle*) debería ser un nombre, puesto que va seguida por un complemento del nombre «*of a model*» y precedida por el determinante «*the*». La quinta omisión (5. *go*) ha de ser un verbo de movimiento precediendo a la preposición «*to*», la octava se reconocerá como parte de la expresión «*of course*», y así sucesivamente.

En algunas ocasiones un hueco puede completarse correctamente con más de una palabra (por ejemplo, también sería posible completar la omisión 2 con «*life*»). Es necesario hacer inferencias extralingüísticas (contexto, cultura, etc.) para buscar la palabra más adecuada y, si varias lo son, queda a juicio del corrector determinar la validez del término según el criterio de corrección que se haya fijado para la prueba.

Mostramos ahora el texto original que corresponde al ejemplo analizado:

Judith Taylor talks about her life as a top model.

I am sure that most people think that the lifestyle of a model is very easy and very exciting. It is true that I go to some fantastic places and I meet some interesting people. And the clothes, of course... I love wearing beautiful clothes!

Aportar información lingüística que no aparece en los mensajes es una actividad normal y relativamente fácil en nuestra lengua materna, pero la dificultad aumenta al intentar recuperar el mensaje en una segunda lengua. A mayor dominio del lenguaje corresponderán mejores resultados, puesto que se será capaz de utilizar mejor la redundancia de la lengua (Spolsky, 1973) y se localizarán mejor las claves lingüísticas y extralingüísticas del texto.

Feldmann y Stemmer (1987) estudiaron la recuperación de las omisiones en el C-test, un tipo de *cloze* (ver apartado 6.5). El proceso que describieron y ahora esbozamos puede aplicarse a cualquier prueba de cierre: la recuperación de cada ítem omitido puede ser automática o no automática. En el primer caso, la palabra surge sin pensar,

mientras que en el segundo el sujeto ha de buscar más claves o estrategias, como las que acabamos de mencionar. Una vez recuperado el término, viene la fase de evaluación, para confirmarlo o rechazarlo. En caso de rechazo, de nuevo comienza la búsqueda o bien llega el abandono. Cuando se duda entre dos términos, la decisión final se suele tomar «por intuición».

3. QUÉ MIDEN LAS PRUEBAS DE CIERRE

Para recomendar la utilización de pruebas de cierre en la enseñanza de lenguas extranjeras, antes hemos de garantizar que con ellas se mide la competencia lingüística del alumno. Lee (1985), Fotos (1991) y Connelly (1997), entre otros, se hacen eco de la controversia acerca de qué miden realmente las pruebas de cierre. Mucho se ha escrito sobre el tema, en este artículo tan sólo citaremos las aportaciones más significativas.

En la literatura encontramos múltiples opiniones; desde los autores que consideran que miden lo mismo que las pruebas de elementos discretos (Farhady, 1979), o que sólo miden las destrezas básicas (Alderson, 1979), hasta los que ven en este tipo de pruebas un instrumento eficaz para medir la competencia lingüística global (Oller, 1979, 1988; Chavez-Oller *et al.* 1985; Bachean, 1982), siendo este último grupo el más numeroso y significativo.

La investigación de Bachman (1982) demostró que la puntuación obtenida en un *cloze* refleja «complex skills ranging along a hierarchy of lower-to higher-order human language processing capacities» y que las pruebas de cierre reflejan un factor de competencia lingüística general «along with three specific traits: a syntactic (clause level) trait, a cohesive (interclausal and intersentential) trait and a strategic (semantic) trait» (Jonz, 1990: 63).

Storey (1997) retomó esta línea de investigación centrándose en las estrategias que cada sujeto emplea en la resolución de las pruebas de cierre. También las investigaciones de Chavez-Oller *et al.* (1985) destacaron que resolver una prueba de cierre implica la puesta en funcionamiento de los mecanismos de «higher-order language processing».

Fotos (1991: 334) planteó que las pruebas de cierre pueden llegar a sustituir a otros instrumentos de evaluación más tradicionales, como los ensayos: «carefully constructed cloze tests have the potential to become useful tools of integrative language assessment

in the EFL situation». También Shohamy (1983) respalda el uso de pruebas de cierre de ratio fija en sustitución de otras pruebas integradoras.

Soyoung Lee (1996) retoma los estudios acerca de la validez concurrente de las pruebas de cierre de ratio fija y su correlación con los ensayos. Entre otros aspectos señala los problemas del ensayo como instrumento de evaluación de la competencia lingüística. A pesar de ser un procedimiento integrador, el ensayo plantea controversias en cuanto al método de corrección (holístico o analítico), el sesgo del tema «writing proficiency may vary with topic», la necesidad de contar con más de un ensayo de cada sujeto para asegurar fiabilidad, etc. (Arthur, 1979; Henning, 1987; Amengual 2003).

A la vista de los resultados, Lee (1996: 62), entre otros, se decanta por las pruebas de cierre como alternativa a otras pruebas: «This result confirms the finding of two previous studies (Fotos, 1991; Hanania and Shikhani, 1986) that cloze tests can be an *alternative* to essay tests» (la cursiva es mía). Además apoya su uso en la práctica docente «as a teaching device in classroom situations».

4. RASGOS FUNDAMENTALES DE LAS PRUEBAS DE CIERRE

Si queremos demostrar que las pruebas de cierre cumplen el objetivo para el cual son diseñadas, hemos de valorar si reúnen las características que debería tener toda prueba de evaluación de la lengua. Validez y fiabilidad se consideran las cualidades básicas de las pruebas, pero no podemos olvidar otros aspectos, como la factibilidad (Bachman y Palmer, 1996).

4.1. Validez y fiabilidad

A pesar de la popularidad de los *clozes*, su validez y fiabilidad han sido cuestionadas a lo largo del tiempo. Puesto que son dos rasgos esenciales, es de vital importancia determinar si las pruebas de cierre los cumplen y en qué grado.

En general, como refleja el apartado anterior, podemos decir que la literatura respalda a las pruebas de cierre como instrumento de medida integrador de la competencia lingüística en EFL (Oller, 1979; Jonz, 1990; Fotos, 1991; Esteban, 2007).

En cuanto a su validez, se ha estudiado la validez de constructo, de contenido, concurrente y predictiva de las pruebas de cierre (Bachman, 1982; Brown, 1983, 1988; Lee, 1985). Según Oller (1979: 347) el hecho de que sean sensibles a «discourse level constraints as well as structural constraints within sentences» es lo que genera coeficientes de validez mayores que los obtenidos con otras pruebas pragmáticas de evaluación.

El estudio de Lee (1996: 69), como los de Fotos (1991) y Hanania y Shikhani (1986), confirma la validez concurrente de los *clozes* como medida de la expresión escrita: «The common integrative nature between the essay and cloze tests is proved».

Hemos de destacar las implicaciones prácticas de la investigación de Lee, pues una vez confirmada la validez de las pruebas de cierre cabe plantearse utilizar la técnica en el aula como «an effective teaching device» más que como mero formato de examen. Así, Hinofotis (1987) y Buck (1988) sugieren, por ejemplo, su uso en ejercicios de comprensión oral, e indican que cada profesor puede hacer las variaciones pertinentes para adaptar las pruebas de cierre a sus propias necesidades educativas en el aula.

Chapelle y Abraham (1990) investigaron la fiabilidad de los distintos tipos de prueba de cierre. Comprobaron que, aunque algunos tipos resultan más fáciles que otros, todos son suficientemente fiables. Demostraron que no hay diferencias notables que afecten a la fiabilidad: «Although statistically significant, these differences in difficulty were not sufficiently large to affect reliability substantially» (1990: 139).

4.2. Factibilidad o carácter práctico

La factibilidad es uno de los aspectos en que encontramos mayor unanimidad en la literatura sobre los *clozes* (Dörnyei y Katona, 1992; Connelly, 1997). Ampliamente se reconoce y alaba su economía de esfuerzo y tiempo. Quizá esta cualidad haya sido básica para impulsar su gran popularidad en el contexto de EFL (*English as a Foreign Language*).

El carácter práctico de las pruebas de cierre radica en su fácil diseño (Oller, 1979). Su creación implica escasas decisiones subjetivas por parte del profesor. Únicamente debe seleccionar el texto base, decidir el tipo de prueba (ratio-fija, variable, C-test, etc.), el punto de comienzo de las omisiones y fijar el criterio de corrección. El profesor gana tiempo en la creación y en la corrección de la prueba, e incluso en la

aplicación, puesto que es aplicable a muchos sujetos a la vez y su administración no requiere un tiempo excesivo.

Debido a sus características, las pruebas de cierre se pueden aplicar repetidamente en el aula como instrumento didáctico en el aprendizaje de la lengua extranjera, y no sólo como pruebas de evaluación (Lee, 1996). El profesor ha de ser quien determine, según su contexto y propósito, cómo utilizar las pruebas de cierre en la enseñanza de lenguas.

5. TIPOS DE PRUEBAS DE CIERRE

Las aportaciones de distintos investigadores interesados en mejorar las pruebas de cierre han dado lugar a varios tipos de *cloze*.

La técnica de cierre no es unitaria (Alderson, 1979: 225). Permite la introducción de variaciones, por ejemplo en el tipo o frecuencia de las omisiones, que dan origen a pruebas bien distintas. Por tanto, el campo de las pruebas de cierre, a pesar del volumen de lo ya investigado, sigue estando hoy abierto a nuevos estudios y aportaciones.

A continuación, ofrecemos una clasificación de las pruebas de cierre atendiendo al tipo y frecuencia de las omisiones. Exponemos también brevemente las ventajas e inconvenientes de cada uno de los tipos de *cloze*.

5.1. De ratio fija

Este tipo de *cloze* se considera el método estándar (Oller, 1979; Chapelle y Abraham, 1990; Jonz, 1990) y es el más utilizado e investigado. Las pruebas de cierre de ratio fija se construyen a partir de un texto previamente seleccionado, en el que se omiten de forma sistemática una cada n palabras. Normalmente el número máximo de omisiones es de 50 (Alderson, 1979; Brown, 1983; Jonz, 1990; Fotos, 1991). Las pruebas resultantes al aplicar una ratio fija a las omisiones se han considerado pruebas pragmáticas muy adecuadas para medir la competencia en una lengua extranjera:

The number of words correctly replaced (by the exact-word scoring procedure) or the number of contextually appropriate words supplied (by the contextually appropriate scoring method) is a kind of overall index of the subject's ability to process the prose in the text (Oller, 1979: 345).

Muchos autores respaldan la técnica incondicionalmente (Oller, 1979; Heilenman, 1983; Hinofotis, 1987; Laesch y Van Kleeck, 1987) por su validez y fiabilidad.

No obstante, el método de ratio fija también ha suscitado dudas en algunos expertos (Alderson, 1979; Klein-Braley, 1983; Brown, 1988) que consideran que, dependiendo del punto de partida de las omisiones y de su frecuencia, un mismo texto puede dar lugar a pruebas de muy distinto grado de dificultad. Oller (1979), por el contrario, insiste en que no afecta donde comiencen las omisiones: «it matters little where the counting begins» (op. cit.: 365) y Bachman (1982) confirma la consistencia de las distintas pruebas de cierre creadas a partir de un mismo texto.

En general, podemos decir que estas pruebas funcionan bien y pueden ser de gran utilidad en la docencia. Corresponde al profesor de Lengua Extranjera decidir la frecuencia de las omisiones. Pero conviene tener en cuenta que si se omite más de la quinta parte de las palabras de un texto (1/5th) muchos de los ítems no pueden ser recuperados, ni siquiera por hablantes nativos (Oller, 1979: 345).

5.2. De ratio variable

Esta variación consiste en crear pruebas de cierre controladas seleccionando las palabras que se van a omitir según un criterio dado, por ejemplo, limitar las omisiones a palabras con carga semántica, o sólo a términos funcionales (Bachean, 1985; Fotos, 1991), dependiendo del propósito de la prueba.

Weir llama a esta técnica *selective deletion gap filling* (en Hughes, 1989: 66) aunque la denominación más habitual es *rational deletion cloze tests*.

Chapelle y Abraham (1990: 124) destacan la mayor fiabilidad y mejor correlación con otras pruebas como ventaja de las pruebas de cierre de ratio variable:

...practically speaking, items selected by experienced text writers may produce tests that are more reliable and more highly correlated with other language tests, especially test measuring traits similar to those particular cloze items were chosen to measure.

Sin embargo, el estudio comparativo de Bachman (1985) entre *clozes* de ratio fija y variable no encontró diferencias significativas entre los resultados de ambos.

Klein-Braley (1997: 62) señala que este tipo de prueba «does enable the test constructor to determine what exactly is being tested», pero también dirige nuestra atención hacia un problema importante: con este sistema de omisiones se pierde el azar, y por tanto se deja de lado uno de los principios de las pruebas de redundancia reducida fijados por Spolsky (1973): «it is not in agreement with the theory of reduced redundancy testing. If the tester chooses what to test, then the random sampling model has been abandoned».

5.3. De elección múltiple

Algunos autores, como Chapelle y Abraham (1990), clasifican a las pruebas de elección múltiple como un tipo de prueba de cierre.

Este método, que fue creado por Jonz (1976), propone al alumno un texto con cierto número de omisiones. Para cada omisión se aportan varias opciones o posibilidades con que completar el texto. Una de ellas es la correcta, y las otras se denominan *distractors*. El alumno debe identificar la respuesta correcta.

Jonz destacó la economía de administración y corrección de la prueba. Pero Klein-Braley (1997: 60) hace la misma crítica que al formato anterior, cuestiona su validez como prueba de redundancia reducida.

Por otra parte, distintas investigaciones muestran un dato que resulta casi obvio: aportar una respuesta propia es más difícil que identificarla (reconocimiento vs. producción).

Weir (1988: 47) señala que este tipo de prueba logra ser fiable al asegurar la objetividad del corrector, pero, sin embargo, las pruebas con este formato deben ser creadas con cuidado para que la pregunta sea clara y para que el alumno no consiga la respuesta correcta simplemente por «eliminación» o «deducción». En consecuencia, su elaboración conlleva mucho tiempo y esfuerzo (Klein-Braley, 1997).

Para Weir (1988), las preguntas de elección múltiple no son un tipo de *cloze*, sino un método diferente de crear exámenes. En su línea, no pensamos que sea adecuado clasificar al test de elección múltiple como tipo de *cloze* porque la tarea propuesta no es sino el reconocimiento o identificación de la respuesta correcta, y por tanto, no mide la competencia en la lengua objeto de estudio. El hecho de que el alumno reconozca y comprenda la palabra no significa que sea capaz de utilizarla en su producción oral o escrita. No obstante, este tipo de prueba puede ser útil en ciertos momentos o situaciones que el profesor ha de determinar.

La popularidad de este tipo de prueba hace que a menudo se identifique la idea de prueba objetiva o test con las pruebas de elección múltiple, como refleja el Diccionario de uso del español de María Moliner (2000).

5.4. Cloze-elide technique

Es otra variación de las pruebas de cierre introducida por Manning (1986, citado en Fotos, 1991). Consiste en insertar palabras incorrectas en un texto, las cuales deben ser detectadas por los alumnos.

Como en el caso de la elección múltiple, creemos que esta técnica no puede ser considerada como prueba de cierre. En primer lugar, porque no presenta al alumno un texto en el que se omite información, sino que ésta ha sido reemplazada por un error. Además, sólo mide el reconocimiento, en este caso del error, pero no la capacidad para producir una respuesta correcta.

5.5. C-test

En los años 80 Klein-Braley y Raatz propusieron un nuevo tipo de prueba de cierre, el C-test. En él las omisiones siguen la *rule of two*, es decir, se elimina la segunda mitad de cada segunda palabra. La tarea consiste en recuperar el texto original completando la segunda mitad de una palabra sí y otra no. Al comienzo y al final del texto se mantiene una parte intacta.

Con esta innovación, Klein-Braley y Raatz pretendían dar a los *clozes* más objetividad y fiabilidad. Y lo lograron (Dörnyei y Katona, 1992; Connelly, 1997; Klein-Braley, 1997; Esteban, 2007). El C-test bien merece un estudio más detallado, pero adelantaremos que, por sus características, podría ser la alternativa a otras pruebas más tradicionales, e incluso formar parte de la actual prueba de Inglés de la Selectividad (Esteban, 2005)³.

6. DISEÑO DE PRUEBAS DE CIERRE

Desde estas páginas animamos a los profesores de Lenguas Extranjeras a conocer mejor la técnica para poder llegar después a crear sus propias pruebas de cierre. El proceso es sencillo; comienza con

³ En *Estudios y criterios para una Selectividad de calidad en el examen de Inglés* (Herrera Soler, H. y J. García Laborda, eds.) aparecen una serie de propuestas para mejorar la actual prueba de Inglés de Selectividad entre la que se incluye la del C-test.

una adecuada selección de textos y culmina con la elección del criterio de corrección. La otra decisión importante es el tipo de prueba de cierre que se desea diseñar, dependiendo de los objetivos que se pretenda conseguir.

6.1. Selección de textos para crear pruebas de cierre

La selección de textos constituye uno de los primeros pasos en la creación de pruebas de cierre. Los expertos insisten en que se haga cuidadosamente (Fotos, 1991), pues de ello puede depender el éxito de todo el proceso.

Oller (1979) asegura que *a priori* cualquier texto puede ser adecuado para crear una prueba de cierre, siempre que tenga la extensión suficiente para efectuar las omisiones⁴. Por supuesto, se ha de tener en cuenta el nivel del alumno en la lengua y el propósito de la prueba. El autor apela al sentido común del profesor. Aconseja evitar los textos que puedan distraer al alumno de su tarea, es decir, aquellos que «involve topics that are intrinsically disturbing or so emotionally charged that they would distract the attention of the students from the main problem set by the test –filling in the blanks» (Oller, 1979: 365) y los que requieran conocimientos técnicos, o contengan temas polémicos (política, religión, aspectos éticos, etc.).

Sasaki (2000: 108) demostró que la inclusión de términos familiares redundaba en una mayor motivación para resolver la prueba. Además facilita al alumno la comprensión del texto y le ayuda a encontrar las claves para completar los ítems. Recomienda elegir textos que sean «familiar enough for the intended examinees to fully use their knowledge».

Teniendo en cuenta estas sencillas indicaciones, cualquier texto puede servir como punto de partida para crear una prueba de cierre.

6.2. Criterios de corrección de las pruebas de cierre

Oller (1979: 367) hace notar que con todos los criterios de corrección de pruebas de cierre investigados se ha constatado la obtención de buenos resultados. Cuando el criterio de corrección es más o menos riguroso lo que se produce es un cambio en la media (*mean score*), pero las variaciones son mínimas: «...there is little change in the relative rank order of scores», pues se mantienen las puntuaciones relativas de los alumnos con respecto al grupo.

⁴ No es conveniente diseñar *clozes* sobre oraciones aisladas: «Cloze items over isolated sentences, of course, do not qualify as pragmatic tests at all and are not recommended» (Oller, 1979: 366).

Brown (1980) valoró los cuatro métodos de corrección que revisaremos en el apartado siguiente: palabra correcta, palabra aceptable, *clozentropy* y formato de elección múltiple. Al igual que Oller, expone que los resultados obtenidos con todos los métodos presentan una alta correlación, por lo tanto, concluye que la elección del método de corrección debe quedar a elección del propio examinador, dependiendo de la situación.

Sin embargo otros estudios (Heilenman, 1983; Hinofotis, 1987; Laesch y Van Kleeck, 1987) atribuyen mayor validez y fiabilidad al método de la palabra exacta.

6.2.1. Palabra exacta

Una de los posibles criterios de corrección para los *clozes* es considerar como válidas exclusivamente las respuestas que se correspondan exactamente con la palabra del texto original. Es también el criterio más estricto y el que menos problemas plantea al corrector, puesto que su tarea se limita a verificar la exactitud incluso ortográfica del término.

6.2.2. Palabra aceptable

Otros autores (Jonz, 1991) respaldan un criterio más razonable y flexible, pero que da mayor cabida a la subjetividad. Consideran correctas las respuestas aceptables para cada omisión y contexto. Proponen el *textually appropriate scoring criterion*. Pero puede surgir el problema de determinar cuáles son concretamente las palabras que se consideran «aceptables» para cada hueco y los límites de la aceptabilidad. Correspondería al profesor o al equipo examinador, en su caso, fijar tales términos como tarea previa a la administración de la prueba.

6.2.3. Clozentropy

Con este criterio se valoran las respuestas comparándolas con las de los hablantes nativos. Este método añade la tarea de administrar previamente cada prueba a un número de hablantes nativos y listar las posibles respuestas. Como hemos indicado para el criterio de la palabra aceptable, la prueba pierde objetividad, aunque quizá gane autenticidad.

6.2.4. Elección múltiple

Si el examinador aporta un repertorio de posibles respuestas para cada omisión la prueba se convierte en un test de elección múltiple.

El alumno sólo debe reconocer la respuesta correcta y señalarla. Su tarea se limita al reconocimiento y no a la producción.

No consideramos adecuado aplicar este criterio de corrección a las pruebas de cierre. La razón es que las convierte en otro tipo de prueba, cuya validez y fiabilidad habría que determinar, pero, en todo caso, distinta en su naturaleza.

Siguiendo las indicaciones de Oller (1979) y Brown (1980) es cada profesor quien debe valorar las posibilidades existentes y elegir un criterio de corrección de *clozes* adecuado para su contexto y objetivos.

7. CONCLUSIONES

Las pruebas de cierre constituyen un procedimiento pragmático de evaluación de la lengua extranjera (Oller, 1979: 42) que permite al profesor valorar la competencia lingüística global, porque reduce la redundancia y obliga al alumno a aplicar la gramática de expectativas de que dispone.

Desde la introducción hemos señalado las múltiples ventajas que aportan las pruebas de cierre, a lo largo del artículo nos hemos centrado en la justificación de su diseño y en los distintos tipos de *cloze* existentes.

En resumen, podemos decir que las pruebas creadas a partir de la técnica de cierre:

- reflejan la competencia lingüística del alumno,
- discriminan su actuación,
- son válidas y fiables como instrumento de medida,
- muy prácticas por su economía de tiempo y esfuerzo.

Su vigencia es también un rasgo destacable, actualmente se sigue investigando sobre pruebas de cierre y siguen apareciendo variaciones que intentan mejorar la técnica (se le puede tomar el pulso en las publicaciones periódicas especializadas en Evaluación de Lenguas Extranjeras, tales como *Language Testing*, *Language Teaching Abstracts*, etc.).

Son inmensas las posibilidades que aporta hoy la técnica de cierre a la enseñanza de lenguas extranjeras, incluso mediante la aplicación de las Nuevas Tecnologías. Con este artículo simplemente hemos intentado acercar un poco más las pruebas de cierre al profesorado de idiomas. Para terminar, animamos a los profesores a iniciarse en

el diseño y aplicación de sus propias pruebas de cierre en las clases, no sólo como instrumento de evaluación sino también como una interesante actividad de aprendizaje.

BIBLIOGRAFÍA

- ALCARAZ, E. y RAMÓN, J. (1980). *La evaluación del inglés. Teoría y práctica*. Madrid: SGEL, S.A.
- ALDERSON, J. C. (1979). «The cloze procedure and proficiency in English as a foreign language». *TESOL Quarterly*, 13, 219-23.
- BACHMAN, L. F. (1985). «Performance on cloze tests with fixed ratio and rational deletions». *TESOL Quarterly*, 19, 535-56.
- BACHMAN, L. F. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- BACHMAN, L. F. y A. S. PALMER (1996). *Language Testing in Practice*. Oxford: Oxford University Press.
- BROWN, J. D. (1983). «A closer look at cloze: Validity and reliability». In J. W. OLLER, Jr. Ed., *Issues in language testing research: 237-50*. Rowley, Massachusetts: Newbury House Publishers.
- BROWN, J. D. (1988). «Tailored cloze: Improved with classical item analysis techniques». *Language Testing* 5, 19-31.
- CHAPELLE, C. A. y R. G. ABRAHAM (1990). «Cloze Method: What Difference does it make?». *Language Testing* 7, 2: 121-146.
- CONNELLY, M. (1997). «Using C-Tests in English with Post-Graduate Students». *English for Specific Purposes* 16 (2), 139-150.
- CULHANE *et al.*, eds. (1984). *Occasional Papers: Practice and problems in language testing*. Colchester: University of Essex.
- DÖRNYEI, Z. y L. KATONA (1992). «Validation of the C-Test amongst Hungarian EFL Learners». *Language Testing* 9, 187-206.
- ESTEBAN GARCÍA, M. (2005). «Niveles de correlación entre el C-test y la prueba de inglés de Selectividad». En HERRERA SOLER, H. y J. GARCÍA LABORDA, eds. *Estudios y criterios para una Selectividad de calidad en el examen de Inglés*, 165-185. Valencia: Ed. Universidad Politécnica de Valencia.
- ESTEBAN GARCÍA, M. (2007). *El C-test: alternativa o complemento de otras pruebas en el aprendizaje del inglés como Lengua Extranjera*. Tesis doctoral. Madrid: Universidad Complutense.

- FELDMANN, U. y B. STEMMER (1987). «Thin__ aloud a__ retrospective da__ in C-te__ taking: diffe__ languages- diff__ learners- sa__ approaches?» En FAERCH, C. y G. KASPER, eds. *Introspection in Second Language Research*. Clevedon: Multilingual Matters Ltd.
- FOTOS, S. (1991). «The Cloze Test as an Integrative Measure of EFL Proficiency: A Substitute for Essays on College Entrance Examinations». *Language Learning*, 41 (3), 313-336.
- JONZ, J. (1991). «Cloze item types and second language comprehension». *Language Testing* 8, 1, 1-22.
- KLEIN-BRALEY, C. (1997). «C-Tests in the context of reduced redundancy testing: an appraisal». *Language Testing* 14, 1: 47-84.
- LEE, S. (1996). «The Concurrent Validity of Cloze Test with Essay Test among Korean Students». *Texas Papers in Foreign Language*, Vol. 2, 2: 57-69.
- OLLER, J. Jr. (1973). «Discrete-point tests versus tests of integrative skills». In OLLER, J. y J. RICHARDS, eds. *Focus on the learner*, 184-199. Rowley, Massachusetts: Newbury House Publishers.
- OLLER, J. Jr. (1973). «Cloze tests of second language proficiency and what they measure». *Language Learning*, 23, 105-18.
- OLLER, J. W. (1979). *Language tests at school*. London: Longman, New York.
- OLLER, J. W., jr. (1983). «Evidence for a general language proficiency factor: an expectancy grammar». In OLLER, J.W., jr. Ed. *Issues in language testing research*, 3-10. Rowley, M.A.: Newbury House Publishers.
- SASAKI, M. (2000). «Effects of cultural schemata on students' test-taking processes for cloze tests: a multiple data source approach». *Language Testing*, 17 (1), 85-114.
- SPOLSKY, B. (1973). «What does it mean to know a language; or how do you get someone to perform his competence?» In OLLER, J. y J. RICHARDS, eds. *Focus on the learner*. Newbury House Publishers; Rowley MA, pp. 164-176.
- TAYLOR, W. L. (1953). «Cloze procedure: -a new tool for measuring readability». *Journalism Quarterly* 30: 414-38.
- WEIR, C. (1988). *Communicative Language Testing*. Exeter Linguistic Studies, Vol. 11. Exeter: University of Exeter.