

## La adaptación de las rúbricas de destreza oral al MCER: Trinity College y EOI

Lucía Fraga Viñas  
*Universidad de A Coruña*  
luciafragav@gmail.com

Fecha de recepción: 15.07.2020

Fecha de aceptación: 12.12.2020

**Resumen:** Uno de los principales objetivos del Marco Común Europeo de Referencia (MCER) es el de servir como base para la elaboración de materiales didácticos, exámenes de certificación, instrumentos de evaluación y currículos. Casi veinte años después de su aprobación, el Marco ha promovido el enfoque comunicativo y ha unificado los niveles de dominio de lenguas a nivel europeo. La intención del presente estudio es la de comprobar cómo se han adaptado las rúbricas de evaluación de la destreza oral en los exámenes del Trinity College y de la Escuela de Idiomas para comprobar si siguen los parámetros establecidos por el Marco.

**Palabras clave:** MCER, rúbricas, evaluación, ILE, exámenes de certificación.

## Adaptation of speaking rubrics to the CEFR: Trinity College and Spain's Official School of Languages

**Abstract:** The Common European Framework of Reference (CEFR) intends to work as a basis for the creation of didactic materials, language certification, instruments of assessment, and curriculums. Almost twenty years after the Framework was passed, the communicative approach has been promoted and the levels of language proficiency have been unified in Europe. This research aims at checking the adaptation of the speaking rubrics of two English Official Certificates: Trinity College and Spain's Official School of Languages in order to determine whether they follow the Framework's guidelines.

**Key words:** CEFR, rubrics, assessment, ESL, English certificates.

**Sumario:** 1. Introducción. 2. La evaluación de la destreza oral y las rúbricas. 3. Metodología. 4. Resultados. 5. Conclusiones.

## 1. Introducción

El campo de la enseñanza y aprendizaje del Inglés como Lengua Extranjera (ILE) en Europa ha experimentado grandes cambios en el siglo XXI, principalmente impulsados por el establecimiento del Marco Común Europeo de Referencia para las Lenguas: Enseñanza, Aprendizaje y Evaluación (MCER). El enfoque educativo que promueve el Consejo Europeo ha revolucionado las metodologías que existían antes de las últimas décadas. La consolidación de la Unión Europea y la necesidad social de adaptarse a las demandas y exigencias del mercado de la globalización ha llevado a centrarse en las destrezas comunicativas. Atrás quedan las clases centradas en la memorización de vocabulario y los ejercicios de gramática, tener fluidez en inglés implica que uno puede mantener una conversación sobre un tema general y entender a cualquier persona sin demasiada dificultad. Por lo tanto, es esencial que la destreza oral y la auditiva se trabajen dentro del aula. Asimismo, la competencia comunicativa, que también implica el dominio de las destrezas de comprensión y producción escritas, requiere que el estudiante tenga conocimientos suficientes para saber qué es adecuado en una comunidad de hablantes en una situación concreta, así como conocimientos sobre pragmática, discurso y cultura o estrategias para solventar aquellas dificultades que pudiesen surgir. En educación, todo esto requiere que la programación, las tareas, la metodología y la evaluación se transformen.

El Marco Común ha proporcionado una fuente de referencia para la creación y el diseño de los currículos educativos y programaciones y además sirve de base para la elaboración de materiales didácticos tales como libros de texto, exámenes de certificación e instrumentos de evaluación. El propio marco indica que “constituye una fuente para el desarrollo de escales de valoración que evalúan el logro de un objetivo concreto de aprendizaje, y los descriptores puede contribuir a la formulación de los criterios” (Consejo de Europa 2002: 179).

A pesar de que en muchos países como Estados Unidos las rúbricas son un instrumento de evaluación común desde principios del siglo XX, en España no se les ha prestado demasiada atención hasta las últimas décadas. En los últimos años, la aparición de rúbricas en los libros de texto se ha hecho más frecuente. No obstante, no se ha realizado demasiada investigación al respecto y aún son muchas las cuestiones por resolver acerca de su aplicación, efectividad, usos y objetividad.

Otra de las finalidades del MCER es la de proporcionar un marco que permita “relacionar sistemas nacionales e institucionales entre sí” y, a la vez, “planear los objetivos de exámenes concretos y de módulos de curso

utilizando las categorías y los niveles de las escalas” (2002: 182). Para ello, ha establecido seis niveles de dominio que van desde usuarios con un dominio básico (A1 y A2), usuarios independientes (B1 y B2) hasta usuarios avanzados (C1 y C2).

| Level            |    |                                   | General description  |
|------------------|----|-----------------------------------|--|
| Proficient User  | C2 | Mastery                           | Highly proficient-can use English very fluently, precisely and sensitively in most contexts. |
|                  | C1 | Effective Operational Proficiency | Able to use English fluently and flexibly in a wide range of contexts.                       |
| Independent User | B2 | Vantage                           | Can use English effectively, with some fluency, in a range of contexts.                      |
|                  | B1 | Threshold                         | Can communicate essential points and ideas in familiar contexts.                             |
| Basic User       | A2 | Waystage                          | Can communicate in English within a limited range of contexts.                               |
|                  | A1 | Breakthrough                      | Can communicate in basic English with help from the listener.                                |

Figura 1. Tabla de niveles del MCER [fuente: English Profile (2015)]

Estos niveles han permitido homogeneizar los exámenes oficiales de certificación de manera que, si un ciudadano cuenta con un nivel, por ejemplo B2, obtenido en España se supone que debe de tener un nivel muy similar a otro que ha obtenido la misma certificación oficial en otro país europeo. De esta manera, no solo los sistemas educativos y los materiales didácticos se han adaptado al MCER y al enfoque comunicativo, también lo han hecho las principales instituciones que otorgan certificados de dominio de lenguas como el Cambridge Assessment English o el Trinity College para inglés, la Alliance Française para francés o el Instituto Cervantes para español como lengua extranjera. Sin embargo, dos décadas después de la implantación del MCER no existen apenas estudios sobre la adecuación o no de la adaptación del marco en los mismos.

Es en este punto donde se enmarca la presente investigación con la intención de analizar dos rúbricas utilizadas como instrumentos de evaluación de la destreza oral en dos exámenes oficiales de certificación. El objetivo es el de comprobar si siguen las directrices establecidas en el marco.

## 2. La evaluación de la destreza oral y las rúbricas

Las destrezas productivas, es decir, aquellas en las que el estudiante debe producir un texto ya sea escrito u oral, se examinan cada vez más con una escala de graduación o rúbrica. No obstante, esto no era así hace unos años. De hecho, la evaluación de la destreza oral es relativamente nueva. No hace mucho tiempo, las metodologías tradicionales como el Método Gramática-Traducción eran dominantes en el aprendizaje de lenguas. En la clase, el alumno apenas hablaba en la lengua meta y se dedicaba a hacer incesantemente ejercicios gramaticales, memorización de listas de vocabulario fuera de contexto y traducción de textos. La situación afortunadamente ha cambiado, principalmente por el impulso del enfoque comunicativo y el establecimiento del MCER.

La introducción de la competencia oral en la clase conlleva cambios en el sistema de evaluación. Examinar esta destreza dista mucho de los métodos de evaluación tradicionales de ejercicios de gramática o vocabulario. En un examen oral no hay respuestas cerradas y hay muchos factores que se deben tener en cuenta, como la manera de articular el mensaje o el lenguaje no verbal. Expertos como Bygate afirman que el proceso oral incluye tres fases principales: la conceptualización del contenido del mensaje, la formulación lingüística del mensaje y la articulación del mensaje (en Baitman & Veliz 2012: 177). Además, la evaluación de la destreza oral en L2 es particularmente compleja por diversos motivos, entre ellos, la dificultad de simular una situación real, el elevado coste de tiempo y recursos que conlleva (Baitman & Veliz 2012: 177).

Son muchos los estudios sobre la evaluación de la destreza oral y los principales problemas que presenta. Referidos específicamente a lenguas extranjeras podemos destacar el de Brittany Baitman y Mauricio Veliz (2012) y el de Emrah Ekmekçi (2016). Mientras el de Ekmekçi, con ochenta estudiantes de inglés como lengua extranjera (ILE) y 6 profesores (tres nativos y tres no nativos), concluyó que las diferencias no eran significativas (2016: 104), el otro, con doce profesores examinando cuatro tareas del TOELF, probó que los examinadores no nativos tienden a poner notas más bajas (2012: 186). Asimismo, los profesores nativos dan más importancia a la fluidez, pronunciación y vocabulario y los no nativos a la precisión gramatical y de vocabulario (2012: 191).

Con respecto a qué criterios examinar en una producción oral, un extenso estudio sobre rúbricas determinó que los aspectos más frecuentes que aparecen en un 95 % de los casos analizados son el rango de vocabulario, el control lingüístico o gramatical, la organización y estructura,

los conectores, el contenido y la pronunciación (Schreiber et al. 2012). En aquellos casos en los que también se examinaba la actuación, el lenguaje no verbal, contacto visual y uso de soportes visuales fueron las dimensiones más habituales.

Es evidente que para evaluar al mismo tiempo todos los criterios mencionados en el párrafo anterior, una rúbrica es el instrumento de evaluación más adecuado. Como ya se ha explicado, la introducción de este instrumento está ligada a la evaluación de la destreza escrita. Antes de su uso, el examinador debía decidir una nota sin ningún tipo de guía ni justificación (Brooks 2012: 229-230). En consecuencia, las notas eran imprecisas y estaban demasiado influenciadas por los errores gramaticales del candidato.

En relación con las investigaciones de rúbricas para evaluar la destreza oral, puede mencionarse un estudio llevado a cabo por Luu Trong Tuan (2012) sobre los resultados obtenidos con rúbricas analíticas y holísticas. El estudio contó con 104 alumnos divididos en dos grupos y examinados con una rúbrica de tipo analítico y otra de tipo holístico respectivamente. La investigación determinó que, después de ser examinados en seis ocasiones y poder ver sus resultados tras cada evaluación en la rúbrica, aquellos que habían sido evaluados con una rúbrica holística apenas mejoraron mientras que aquellos evaluados con una rúbrica analítica mejoraron su calificación en más de un punto. Esto sugiere que las rúbricas de tipo analítico ayudan a los estudiantes a mejorar más que las de tipo holístico.

### 3. Metodología

El Consejo de Europa estableció ya en 1971 que la enseñanza se ha de planificar como un conjunto coherente, cubriendo las especificaciones de objetivos, la evaluación de logros y la efectividad del sistema y que eso implica un esfuerzo coordinado por parte de las administraciones educativas, materiales, examinadores, inspectores, profesores y estudiantes, que siempre deben compartir objetivos y criterios de evaluación (Vez 2011: 89).

Esta investigación utilizó un enfoque cualitativo, que es el que no usa datos numéricos extraídos de la realidad, sino que interpreta y describe la realidad en detalle a través de palabras (Madrid 2001: 12). Además, la principal técnica de investigación fue el análisis de documentos. Esta técnica permite formular inferencias válidas que puede aplicarse a un contexto a partir de ciertos datos (Krippendorff en Gil 2011:283).

Con el objetivo de aplicar el método cualitativo de manera rigurosa, se siguieron los siete pasos establecidos por Gil Pascual. El primero de ellos consistió en establecer los objetivos y el contexto. En este punto se determinaron los objetivos de la investigación que fueron los de analizar las rúbricas utilizadas por dos certificados oficiales de inglés para comprobar su adecuación a las directrices marcadas por el MCER; y determinar qué tipos de rúbricas son más comunes para poder establecer patrones.

El segundo paso fue la definición de las unidades de análisis. En este paso se seleccionaron las rúbricas de los certificados que iban a ser estudiadas. Evidentemente, existe una amplia gama de certificaciones oficiales de inglés como lengua extranjera. Debido a que es objeto del estudio determinar si siguen las directrices del MCER, solo se tuvieron en cuenta aquellos que perteneciesen a instituciones europeas. Las instituciones seleccionadas fueron el Trinity College y la Escuela Oficial de Idiomas. El nivel seleccionado fue el de B2, que corresponde con el intermedio alto y es el que se pide de forma más habitual para trabajar y estudiar en Europa. En este nivel y, según el MCER, un usuario puede interactuar con cierto grado de fluidez y espontaneidad que hace la interacción regular con un nativo bastante posible, sin gran dificultad para ninguna de las dos partes (Consejo de Europa, 2002: 24).

| Score    | Communicative effectiveness   | Interactive listening   | Language control  | Delivery  |
|----------|---|---|---|---|
|          | <ul style="list-style-type: none"> <li>▫ Task fulfilment</li> <li>▫ Appropriacy of contributions/turn-taking</li> <li>▫ Repair strategies</li> </ul>  | <ul style="list-style-type: none"> <li>▫ Comprehension and relevant response</li> <li>▫ Level of understanding</li> <li>▫ Speech rate of examiner interventions</li> <li>▫ Speed and accuracy of response</li> </ul>  | <ul style="list-style-type: none"> <li>▫ Range</li> <li>▫ Accuracy/precision</li> <li>▫ Effects of inaccuracies</li> </ul>  | <ul style="list-style-type: none"> <li>▫ Intelligibility</li> <li>▫ Lexical stress/intonation</li> <li>▫ Fluency</li> <li>▫ Effects on the listener</li> </ul>  |
| <b>4</b> | <ul style="list-style-type: none"> <li>▫ Fulfills the task very well</li> <li>▫ Initiates and responds with effective turn-taking</li> <li>▫ Effectively maintains and develops the interaction</li> <li>▫ Solves communication problems naturally, if any</li> </ul>   | <ul style="list-style-type: none"> <li>▫ Understands all interventions on a first hearing</li> <li>▫ Interprets examiner's aims and viewpoints accurately by making links with earlier information</li> <li>▫ Makes immediate and relevant responses</li> </ul> | <ul style="list-style-type: none"> <li>▫ Uses a wide range of grammatical structures/lexis flexibly to deal with topics at this level</li> <li>▫ Consistently shows a high level of grammatical accuracy and lexical precision</li> <li>▫ Errors do not impede communication</li> </ul>   | <ul style="list-style-type: none"> <li>▫ Clearly intelligible</li> <li>▫ Uses focal stress and intonation effectively</li> <li>▫ Speaks promptly and fluently</li> <li>▫ Requires no careful listening</li> </ul>   |
| <b>3</b> | <ul style="list-style-type: none"> <li>▫ Fulfills the task appropriately</li> <li>▫ Initiates and responds appropriately</li> <li>▫ Maintains and develops the interaction appropriately (eg expanding and developing ideas, and showing understanding of what the examiner said)</li> <li>▫ Deals with communication problems well</li> </ul>  | <ul style="list-style-type: none"> <li>▫ Understands most interventions on a first hearing</li> <li>▫ Interprets examiner's aims and viewpoints accurately</li> <li>▫ Makes prompt and relevant response</li> </ul>   | <ul style="list-style-type: none"> <li>▫ Uses an appropriate range of grammatical structures/lexis to deal with topics at this level</li> <li>▫ Shows a relatively high level of grammatical accuracy and lexical precision</li> <li>▫ Errors do not impede communication</li> </ul>  | <ul style="list-style-type: none"> <li>▫ Clearly intelligible despite some use of non-standard phonemes</li> <li>▫ Uses focal stress and intonation appropriately</li> <li>▫ Generally speaks promptly and fluently – occasionally affected by some hesitancy</li> <li>▫ Requires almost no careful listening</li> </ul>                                |
| <b>2</b> | <ul style="list-style-type: none"> <li>▫ Fulfills the task acceptably with support</li> <li>▫ Initiates and responds acceptably</li> <li>▫ Maintains and develops the interaction, but contributions are not always appropriate and/or somewhat dependent on the examiner</li> <li>▫ Manages to solve communication problems, but requires more than one attempt and/or does not always do this naturally (eg 'What?')</li> </ul> | <ul style="list-style-type: none"> <li>▫ Usually understands interventions; occasionally needs clarification</li> <li>▫ Shows occasional uncertainty about examiner's aims or viewpoints</li> <li>▫ Makes relatively prompt responses</li> </ul>                | <ul style="list-style-type: none"> <li>▫ Uses an acceptable range of grammatical structures/lexis to manage topics at this level, but grammatical/lexical gaps still cause hesitation and circumlocution</li> <li>▫ Shows an acceptable level of grammatical accuracy and lexical precision</li> <li>▫ Most errors do not impede communication</li> </ul> | <ul style="list-style-type: none"> <li>▫ Intelligible despite some use of non-standard phonemes</li> <li>▫ Uses focal stress and intonation acceptably</li> <li>▫ Speaks promptly and fluently enough to follow</li> <li>▫ Requires some careful listening</li> </ul>   |
| <b>1</b> | <ul style="list-style-type: none"> <li>▫ Does not fulfil the task even with support</li> <li>▫ Does not initiate or respond adequately</li> <li>▫ Does not maintain and develop the interaction sufficiently</li> <li>▫ Contributions are inappropriate and/or overly dependent on the examiner</li> <li>▫ Has some difficulty in resolving communication problems</li> </ul>   | <ul style="list-style-type: none"> <li>▫ Has difficulty in understanding interventions</li> <li>▫ Frequently misinterprets examiner's aims and viewpoints</li> <li>▫ Responds slowly due to difficulty in understanding input</li> </ul>                        | <ul style="list-style-type: none"> <li>▫ Uses a limited range of grammatical structures/lexis that is not always adequate to deal with topics at this level</li> <li>▫ Does not show an adequate level of grammatical accuracy and lexical precision</li> <li>▫ Some errors impede communication</li> </ul>   | <ul style="list-style-type: none"> <li>▫ Generally intelligible or sometimes unintelligible. Use of non-standard phonemes is sometimes or frequently evident</li> <li>▫ Sometimes or often misuses focal stress and intonation</li> <li>▫ Speaks slowly. Sometimes or often halted by hesitancy</li> <li>▫ Requires (some) careful listening</li> </ul> |
| <b>0</b> | No performance to assess (candidate does not speak, or does not speak in English). Also use if no topic is prepared.  |   |   |   |

Figura 2. Rúbrica para evaluar la destreza oral del Trinity College ISE-II

**CRITERIOS DE VALORACIÓN DE LA EXPRESIÓN E INTERACCIÓN ESCRITA DE NIVEL AVANZADO**

| GRADO DE ADECUACIÓN AL NIVEL EVALUADO  | Nada adecuado al nivel (0 puntos)  | Muy poco adecuado al nivel (1 punto)  | Poco adecuado al nivel (2 puntos)   | Aceptable para el nivel (3 puntos)  | Destacado para el nivel (4 puntos)  | Excelente para el nivel (5 puntos)  |
|--|--|---|---|---|---|---|
| <b>ADECUACIÓN ORTOGRÁFICA Y DE PUNTUACIÓN</b>  | Comete graves y sistemáticos errores ortográficos. Hay ausencia y/o uso arbitrario de los signos de puntuación.  | Comete muchos errores ortográficos y/o de puntuación.   | Comete errores ortográficos y/o de puntuación.  | Comete pocos errores ortográficos y/o de puntuación.  | Comete algún error esporádico de ortografía y/o de puntuación.  | Utiliza con corrección las convenciones ortográficas y de puntuación.   |
| <b>ADECUACIÓN GRAMATICAL</b>   | Las estructuras gramaticales utilizadas no se adecúan a la situación comunicativa y/o comete sistemáticamente errores gramaticales.  | Las estructuras gramaticales utilizadas se adecúan muy poco a la situación comunicativa y/o comete muchos errores gramaticales, comunicándose de forma muy parcial.   | Las estructuras gramaticales utilizadas no se adecúan siempre a la situación comunicativa y/o comete errores gramaticales frecuentes, comunicándose de forma parcial.   | Las estructuras gramaticales utilizadas se adecúan prácticamente siempre a la situación comunicativa y comete pocos errores gramaticales, comunicándose de forma aceptable.   | Las estructuras gramaticales utilizadas se adecúan a la situación comunicativa, aunque comete errores gramaticales esporádicos de carácter leve, comunicándose correctamente.   | Utiliza con corrección las estructuras gramaticales del nivel, comunicándose de forma excelente.  |
| <b>ADECUACIÓN LÉXICA</b>   | El vocabulario utilizado no se adecúa a la situación comunicativa y/o comete sistemáticamente errores léxicos.   | El vocabulario utilizado se adecúa muy poco a la situación comunicativa y/o comete muchos errores léxicos, comunicándose de forma muy parcial.  | El vocabulario utilizado no se adecúa siempre a la situación comunicativa y/o comete errores léxicos frecuentes, comunicándose de forma parcial.  | El vocabulario utilizado se adecúa a la situación comunicativa prácticamente siempre y comete pocos errores léxicos, comunicándose de forma aceptable.  | El vocabulario utilizado se adecúa a la situación comunicativa, aunque comete errores léxicos esporádicos de carácter leve, comunicándose correctamente.  | Utiliza con corrección el vocabulario del nivel comunicándose de forma excelente.   |
| <b>ADECUACIÓN DEL NIVEL DE FORMALIDAD Y/O RESPETO AL PROPOSITO COMUNICATIVO, A LA SITUACIÓN Y/O AL RECEPTOR.</b>   | No ajusta el nivel de formalidad y/o respeto al propósito comunicativo, a la situación y/o al receptor.  | Apenas ajusta el nivel de formalidad y/o respeto al propósito comunicativo, a la situación y/o al receptor.   | Ajusta de manera insuficiente el nivel de formalidad y/o respeto al propósito comunicativo, a la situación y/o al receptor.   | Ajusta suficientemente el nivel de formalidad y/o respeto al propósito comunicativo, a la situación y/o al receptor.  | Ajusta sin dificultad aparente y con acierto el nivel de formalidad y/o respeto al propósito comunicativo, a la situación y/o al receptor.  | Ajusta con precisión el nivel de formalidad y/o respeto al propósito comunicativo, a la situación y/o al receptor.  |
| <b>ORGANIZACIÓN DEL DISCURSO: CARACTERÍSTICAS PARA CREAR UNA ESTRUCTURA LÓGICA, ORGANIZADO Y DISTRIBUYENDO EL TEXTO CONFORME AL FORMATO Y TIPO DE TEXTO REQUERIDO Y UTILIZANDO MECANISMOS DE COHESIÓN ADECUADOS.</b>       | Elabora un discurso incoherente y/o nada claro que no se ajusta al tipo de texto requerido ni a las convenciones de organización y distribución de párrafos. No utiliza mecanismos de cohesión, palabras ni frases clave.  | Elabora un discurso muy poco coherente y/o claro que apenas se ajusta al tipo de texto requerido y/o a las convenciones de organización y distribución de párrafos. Apenas utiliza o utiliza incorrectamente mecanismos de cohesión, palabras y frases clave. | Elabora un discurso poco coherente y/o claro que se ajusta sólo parcialmente al tipo de texto requerido y/o a las convenciones de organización y distribución de párrafos. No siempre utiliza con corrección mecanismos de cohesión, palabras y frases clave. | Elabora un discurso suficientemente coherente y/o claro que se ajusta al tipo de texto requerido y a las convenciones de organización y distribución de párrafos. Utiliza con suficiente corrección mecanismos de cohesión, palabras y frases clave, aunque comete algunos errores.           | Elabora un discurso muy coherente y/o claro que se ajusta al tipo de texto requerido y a las convenciones de organización y distribución de párrafos. Utiliza con corrección mecanismos de cohesión, palabras y frases clave, aunque haya esporádicamente alguna incoherencia apenas relevante. | Realiza un discurso totalmente coherente y/o claro, que se ajusta al tipo de texto requerido y a las convenciones de organización y distribución de párrafos. Utiliza con total corrección mecanismos de cohesión, palabras y frases clave.             |
| <b>INTERACCIÓN: CAPACIDAD PARA INICIAR, MANTENER Y TERMINAR EL DISCURSO. CAPACIDAD PARA INTERACTUAR CON FLUIDEZ Y NATURALIDAD.</b>   | No es capaz de intercambiar, registrar y/o comentar información de forma pertinente por lo que no se ajusta a las expectativas del receptor. No utiliza expresiones para iniciar, mantener y dar por terminado el discurso. Escribe en un estilo totalmente artificial y rígido. | Intercambia, requiere y/o comenta información ajustándose muy poco a las expectativas del receptor. Apenas utiliza o utiliza incorrectamente expresiones para iniciar, mantener y dar por terminado el discurso. Escribe con muy poca fluidez y naturalidad.  | Intercambia, requiere y/o comenta información ajustándose poco a las expectativas del receptor. No siempre utiliza con corrección expresiones para iniciar, mantener y dar por terminado el discurso. Escribe con poca fluidez y naturalidad.                 | Intercambia, requiere y/o comenta información ajustándose suficientemente a las expectativas del receptor. Utiliza con suficiente corrección expresiones para iniciar, mantener y dar por terminado el discurso, aunque comete algunos errores. Escribe con suficiente fluidez y naturalidad. | Intercambia, requiere y/o comenta información respondiendo con mucho acierto a las expectativas del receptor. Utiliza con corrección expresiones para iniciar, mantener y dar por terminado el discurso, aunque comete algún error esporádico. Escribe con mucha fluidez y naturalidad.         | Intercambia, requiere y/o comenta información respondiendo con total acierto a las expectativas del receptor. Utiliza con total corrección expresiones para iniciar, mantener y dar por terminado el discurso. Escribe con total fluidez y naturalidad. |
| <b>DESARROLLO DEL TEMA: CAPACIDAD PARA DESCRIBIR Y PRESENTAR SITUACIONES, NECESIDADES, HECHOS Y OPINIONES, NARRAR, EJEMPLIFICAR Y ARGUMENTAR DE FORMA CLARA Y PERTINENTE AL TEMA PROPUESTO, CON FLUIDEZ Y NATURALIDAD.</b> | No es capaz de describir, presentar situaciones, necesidades, hechos y/o opiniones, narrar, explicar y/o argumentar. Escribe en un estilo totalmente artificial y rígido.  | Describe, presenta situaciones, necesidades, hechos y/o opiniones, narra, explica y/o argumenta con mucha dificultad. Escribe con muy poca fluidez y naturalidad.   | Describe, presenta situaciones, necesidades, hechos y/o opiniones, narra, explica y/o argumenta con dificultad. Escribe con poca fluidez y naturalidad.   | Describe, presenta situaciones, necesidades, hechos y/o opiniones, narra, explica y/o argumenta de manera adecuada, sin mucha dificultad. Escribe con suficiente fluidez y naturalidad.   | Describe, presenta situaciones, necesidades, hechos y/o opiniones, narra, explica y/o argumenta correctamente, sin apenas dificultad. Escribe con mucha fluidez y naturalidad.  | Describe, presenta situaciones, necesidades, hechos y/o opiniones, narra, explica y/o argumenta de forma excelente, sin dificultad aparente. Escribe con total fluidez y naturalidad.   |

PRUEBAS TERMINALES ESPECÍFICAS DE CERTIFICACIÓN DE IDIOMAS  
Escuela Oficial de Idiomas del Principado de Asturias

Figura 3. Rúbrica para evaluar la destreza oral de la EOI Nivel Avanzado (B2)

El tercer paso fue el de determinar las reglas de numeración y recuento, es decir, definir los principios que se iban a usar para determinar qué rúbricas siguen las directrices del marco. En este caso, se acudió al MCER que establece que las rúbricas deben de ser instrumentos factibles por lo que no deben de medir más de cinco criterios. En referencia a los descriptores, se detalla que deben estar escritos en positivo, ser breves pero precisos, evitando cualquier vaguedad en su formulación. Además, también se clasificaron las rúbricas según su tipología para saber qué tipo de rúbrica era en cuanto a su manera de medir, a la escala utilizada, su tema, su nivel de aplicación, su función, el tipo de examinador y formato.

La categorización y codificación forman el cuarto paso. El instrumento escogido fueron las tablas de recogida de datos. Estos cuadros se diseñaron específicamente para el estudio y teniendo en cuenta que la información que iban a contener iba a ser examinada y comparada y, naturalmente, analizada. La construcción visual de las tablas fue muy útil para categorizar los datos extraídos de las rúbricas.

|          |             |          |  |
|----------|-------------|----------|--|
| TYPE     | measurement |          |  |
|          | scoring     |          |  |
|          | theme       |          |  |
|          | application |          |  |
|          | function    |          |  |
|          | scorer      |          |  |
|          | channel     |          |  |
| RELEVANT |             |          |  |
| VALID    |             |          |  |
| CEFR     | scoring     |          |  |
|          | descriptors | Pos.     |  |
|          |             | Brief.   |  |
|          |             | no vague |  |

Figura 4. Tabla de recogida de datos

A continuación, Gil recomienda medir la fiabilidad y la validez. Según Phelan and Wren (2005-2006, párrafo 2) la validez de una rúbrica viene determinada por cuán bien mide lo que tiene que examinar. Con respecto a la fiabilidad, esta se refiere al grado en el que un instrumento produce resultados estables y consistentes, normalmente con un coeficiente como el Alpha Cronbach.

Por último, los datos se valoraron en su conjunto y se compararon con los objetivos de detectar incongruencias con el MCER y establecer patrones de similitudes o diferencias entre las rúbricas.

#### 4. Resultados

Después de un análisis comparatista de las rúbricas utilizadas para examinar la destreza oral en el Trinity College ISE-II y en el nivel avanzado de la EOI (B2), se encontraron coincidencias tales como el hecho de que ambas rúbricas son de tipo analítico y no holístico. Ya hemos mencionado anteriormente que investigaciones como las de Trong (2012) prueban que las rúbricas de este tipo son más fiables. Otros estudios también llegaron a las mismas conclusiones (Sundeen, 2014; Becker, 2016).

Ambas rúbricas utilizan una escala numérica, de cero a cuatro puntos en el caso del ISE-II y de cero a cinco en la EOI. La principal diferencia es que la rúbrica empleada para examinar el B2 de la EOI contiene una escala cualitativa además de la numérica que consiste en expresiones tales como “nada adecuado al nivel”, “aceptable para el nivel”, “excelente para el nivel”,

etc. En general, todas las rúbricas que se usan en certificados oficiales contienen una escala numérica en un ejercicio de precisión, aunque el empleo de una escala cualitativa, si se acompaña de un comentario para el alumno, puede aportar al alumno una retroalimentación más precisa de cómo ha sido su actuación.

En cuanto a su aplicación, la rúbrica de la Escuela Oficial de Idiomas es una rúbrica de destreza, esto quiere decir que solo se usa para examinar una destreza, en este caso la oral, pero no para todas sus tareas. En el caso de la rúbrica del Trinity College, solo se utiliza una rúbrica para evaluar la destreza oral, pero tiene la peculiaridad de que dicha rúbrica evalúa, al mismo tiempo, la destreza auditiva. Esto ocurre porque ambos exámenes se realizan de manera conjunta y sus tareas son diferentes de las habituales en un examen de destreza oral. Mientras que en la mayoría de los exámenes orales de certificación el candidato debe elaborar un monólogo, describir una foto, o interactuar en una discusión con un compañero, en el del Trinity College el estudiante deberá escuchar un audio y después resumirlo y contarlo con sus palabras, de manera que quede patente que lo ha entendido. A continuación, el examinador y el candidato establecerán una conversación con preguntas y respuestas sobre temas relacionados con la prueba auditiva.

El contraste de las rúbricas con las directrices que contiene el MCER ha arrojado resultados interesantes. Como ya se ha indicado, el marco propone el uso de instrumentos que sean realmente factibles. Para ello, una escala de evaluación debe tener entre cuatro y cinco criterios, ser cómoda y manejable, lo cual implica que se presente en una hoja con descriptores que no sean excesivamente largos puesto que debe examinar con la misma rúbrica a un número elevado de estudiantes.

La rúbrica empleada para medir la destreza oral del Trinity College ISE-II se presenta en una sola hoja con cuatro criterios: eficacia comunicativa, destreza auditiva interactiva, control del lenguaje y actuación, por lo que, grosso modo, podría ser una rúbrica factible o manejable. Por el contrario, la rúbrica utilizada en la EOI mide hasta siete criterios diferentes: adecuación ortográfica y de puntuación, adecuación gramatical, adecuación léxica, adecuación del discurso, organización del discurso, interacción y desarrollo del tema. Además, dos de ellos se subdividen en dos: adecuación gramatical en corrección de las estructuras y riqueza y variedad de las estructuras; y adecuación léxica en corrección del vocabulario y riqueza y variedad del mismo. Es, por lo tanto, una rúbrica no factible o manejable, examinar tantos criterios consume demasiado tiempo por alumno y no sería adecuada para evaluar a un gran número de candidatos en un período razonable de tiempo.

Como ya hemos comentado, el marco también pauta cómo deben ser los descriptores: deben estar escritos con palabras positivas y ser breves, pero a la vez precisos. Los resultados de la investigación llevada a cabo señalan que en ninguna de las rúbricas analizadas los descriptores están escritos únicamente en positivo pese a que en ambos casos se aprecian ciertos intentos por mantener un lenguaje positivo. Sin embargo, en los niveles más bajos de la escala de la EOI encontramos expresiones como “comete muchos errores”, “comete graves y sistemáticos errores”, “comete muchos errores léxicos”, “no es capaz de describir situaciones”. En la rúbrica del ISE-II, por su parte, hay frases como “no responde adecuadamente”, “no mantiene la interacción suficientemente”, si bien se aprecia un intento por utilizar formulaciones positivas “habla despacio”, “usa un número limitado de estructuras gramaticales”, etc.

Los descriptores analizados no son vagos o imprecisos, todo lo contrario, son muy detallados, aunque esto hace que sean demasiado largos y con ello que el tiempo de evaluación sea excesivo.

## 5. Conclusiones

La investigación realizada arroja unos resultados interesantes sobre los que se puede reflexionar. El estudio realizado tenía como principal objetivo determinar el nivel de adecuación de los instrumentos de evaluación de certificados oficiales de inglés al Marco Común.

En primer lugar, cabe señalar que las rúbricas no siempre están disponibles para los candidatos de los certificados oficiales de inglés. En el caso de la Escuela Oficial de Idiomas, depende de cada escuela el hacerla pública o no y no todas lo hacen. En el caso del Trinity College, sus rúbricas de evaluación sí están disponibles en la página web de la institución. Los alumnos que disponen de acceso a la rúbrica antes de un examen o a los que se les enseña a utilizar la rúbrica con la que van a ser evaluados tienden a mejorar sus resultados según diversos estudios como, por ejemplo, el llevado a cabo por Laurian & FitzGerald en 2013. Por eso, sería recomendable que todas las instituciones fuesen transparentes en esta área y proporcionasen, a través de su página web, un acceso claro a las rúbricas de evaluación que van a utilizar. Además, sería interesante que también realizasen estudios sobre la fiabilidad y validez de sus certificados con la intención de mejorarlos y que compartiesen los resultados de estos por medio de la publicación de coeficientes como el Alpha Cronbach.

En segundo lugar, es importante destacar el hecho de que ninguna de las escalas de graduación analizadas sigue completamente las

directrices marcadas en el MCER. La rúbrica del Trinity College ISE-II sí presenta cuatro criterios como dicta el Marco, pero la de la EOI tiene ocho, lo cual ralentiza enormemente el proceso de evaluación por el consumo de tiempo que implica evaluar a cada candidato con ella. Los descriptores de ambas rúbricas no se ajustan a lo esperado de ellos porque no siempre están escritos en positivo y no son breves. No obstante, ninguna de las rúbricas presenta descriptores vagos o poco precisos, al contrario, los descriptores son claros pero muy extensos.

El hecho de que ambas rúbricas coincidan en no utilizar siempre descriptores totalmente positivos, particularmente en los niveles inferiores de la escala, tal y como dicta el MCER, pone en cuestión si esto es realmente factible o no. Esto es debido a que ser capaz de describir de manera concisa una actuación negativa utilizando palabras positivas es realmente complicado.

Otra desventaja que ofrece el MCER tiene que ver con que todas las escalas que se incluyen en el mismo salvo una son de tipo holístico. En consecuencia, son las propias instituciones las que deben crear sus propias escalas a partir del Marco. Pongamos por caso que se quiere crear una rúbrica para la destreza escrita del nivel B2. En este caso, habría que ir a la escala de tipo holístico para la destreza escrita del marco. En dicha escala, cada nivel se corresponde con uno de los niveles del marco, desde A1 hasta C2, por lo que habría que seleccionar el nivel B2 y a partir de ese único descriptor crear una rúbrica nueva con sus criterios y niveles. Esto provoca que las rúbricas se distancien mucho del Marco y sean muy diferentes y diversas entre sí y, con ello, no siempre válidas, efectivas o fiables. Si dentro del MCER se incluyesen escalas analíticas para todas las destrezas, este proceso se agilizaría. Evidentemente, las rúbricas creadas a partir de ellas podrían adaptarse a cada examen, tarea o circunstancias, pero aun así serían más similares en cuanto a criterios o formulación de descriptores. La inclusión de escalas de tipo analítico también propiciaría más ejemplos de descriptores en positivo, breves y precisos.

En una sociedad cada vez más globalizada y concretamente en un marco europeo que promueve la movilidad de sus ciudadanos para trabajar, estudiar y viajar, el Consejo de Europa ha llevado a cabo una unificación de niveles de dominio esencial. El MCER es además extremadamente útil en cuanto a la evaluación. Sin embargo, su aplicación y adaptación no está tan bien implantada como debería y por ello es necesario que se revise y se mejore para que se ajuste a la realidad educativa y a las nuevas realidades o problemas que puedan surgir. El MCER, al igual que cualquier lengua, debe evolucionar y desarrollarse constantemente.

## Referencias bibliográficas

- BAITMAN, B. and VELIZ, M. (2012): "A Comparison of oral evaluation ratings by native English teachers and non-native English speaker teachers". *Literatura y Lingüística*, 61 (3), 171-200.
- BECKER, A. (2016): "Student-generated scoring rubrics: Examining their formative value for improving ESL students' writing performance". *Assessing Writing*, volume 29, 15-24.
- BROOKS, G. (2012): "Assessment and Academic Writing: A look at the Use of Rubrics in the Second Language Writing Classroom". *Kwansei Gakuin University Humanities Review*, 17, 227-240.
- CONSEJO DE EUROPA (2002): *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Instituto Cervantes y Anaya. Disponible en web: <[https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/marco/](https://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/)> [Fecha de última consulta: 27 de junio de 2020]
- EKMEKÇI, E. (2016): "Comparison of Native and Non-native English language Teacher's Evaluation of EFL Learner's Speaking Skills: Conflicting or Identical Rating Behaviour?" *English Language Teaching*, 9, (5), 98-105.
- ESCUELA OFICIAL DE IDIOMAS DE GIJÓN. (2017a): Criterios, procedimientos e instrumentos de evaluación, *eoigijon*. Disponible en web: <<http://eoigijon.com/idiomas/ingles/>> [Fecha de última consulta: 15 de mayo de 2018]
- ESCUELA OFICIAL DE IDIOMAS DE GIJÓN. (2017b): Modelos de Pruebas de Certificación de idiomas. Inglés. Nivel Avanzado (NA). *eoigijon*. Disponible en web: <<https://www.educastur.es/estudiantes/idiomas/pruebas-certificacion/modelos>> [Fecha de última consulta: 15 de mayo de 2020]
- GIL PASCUAL, J.A. (2011): *Técnicas e instrumentos para la recogida de información*. Madrid: Universidad Nacional de Educación a Distancia
- LAURIAN, S. and FITZGERALD, C. J (2013): "Effects of using rubrics in a university academic level Romanian literature class". *Procedia. Social and Behavioral Sciences*, Elsevier, 76, 431-440.
- MADRID, D. (2001): "Introducción a la investigación en el aula de la lengua extranjera". *Metodología de investigación en el área de filología inglesa*, editado por María Elena García Sánchez y María Sagrario Salaberri. Almería: Universidad de Almería, 11-45.
- PHELAN, C. and WREN, J. (2005-2006): *Exploring reliability in academic assessment*, Iowa: UNI Office of Academic Assessment, University of Northern Iowa.

- PROFILE, E. (2015). *English Profile: The CEFR for English*. Cambridge, UK: Cambridge University Press. Disponible en web: <<https://www.englishprofile.org/the-cefr/cefr-for-teachers-learners>> [Fecha de última consulta: 14 de mayo de 2020]
- SCHREIBER, L. M. et al. (2012): "The Development and Test of the Public Speaking Competence Rubric". *Communication Education*, Routledge, 61 (3), 205-233
- SUNDEEN, T. H. (2014): "Instructional rubrics: Effects of presentation on writing quality". *Assessing writing*, volume 21, 74-87.
- TRINITY COLLEGE LONDON (2017): Integrated Skills in English (ISE) Guide for Teachers — ISE-II (B2). *Trinity College* web page, Online edition Jun. Disponible en web: <<https://www.trinitycollege.com/qualifications/english-language/ISE/ISE-II-B2-resources/ISE-II-B2-Guides>> [Fecha de última consulta: 18 de mayo de 2020]
- TRONG TUAN, L. (2012): "Teaching and Assessing Speaking Performance through Analytic Scoring Approach". *Theory and Practice in Language Studies, Academy Publisher*, 2 (4) 673-679,
- VEZ, J. M. (2011): "La Investigación en Didáctica de las Lenguas Extranjeras". *Educatio Siglo XXI*, 29 (1) 81-108.