

# ESPECIFICACIONES PARA UNA PRUEBA DE COMPREENSIÓN AUDITIVA Y DE EXPRESIÓN E INTERACCIÓN ORALES EN UN ENTORNO ESCOLAR

JOSÉ CUADRADO MORENO

MARÍA REYES FERNÁNDEZ

MANUEL GREGORIO CABALLERO CALAVIA

UNIVERSIDAD PABLO DE OLAVIDE (SEVILLA, ESPAÑA)

**Resumen:** Se presenta un procedimiento para desarrollar la estructura de las subpruebas de comprensión auditiva y de expresión e interacción orales de una prueba de aprovechamiento estandarizada de inglés para el alumnado del segundo curso de E.S.O. Sevilla (España). Este procedimiento está basado en un estudio observacional de las actividades de aprendizaje en las clases de inglés del segundo curso de Educación Secundaria Obligatoria en Sevilla, en una revisión del currículo de inglés para este curso y en pruebas similares de lengua extranjera. Al contrario de los métodos tradicionales de recoger evidencias basadas en el contenido para las pruebas educativas, este procedimiento se realiza antes de escribir los ítems y las tareas de la prueba.

**Palabras claves:** Validez del contenido, prueba de inglés, primer ciclo de educación secundaria, prueba estandarizada.

**Abstract:** A procedure is presented to develop the structure of the subtests of oral reception (listening) and oral production and interaction (speaking) in a standardised achievement test of English for Year-8 pupils in Sevilla (Spain). This procedure is based on an observational study of the learning activities in the Year-8 classes of English in Sevilla and on a review of both the Year-8 curriculum of English and similar foreign language tests. As opposed to the usual ways of collecting content-based evidence for educational tests, this procedure is carried out before both test items and tasks are written.

**Key words:** Content validity, test of English, lower secondary education, standardized test.

**Résumé:** On présente ici une méthode pour développer la structure des différentes parties – compréhension à l'audition, expression orale et interaction- d'un examen en langue anglaise aux élèves de 2<sup>ème</sup> année de la ESO à Séville (Espagne). Cette méthode est basée sur une étude observationnelle des activités d'apprentissage utilisées au cours d'anglais en deuxième année de l'Enseignement Secondaire Obligatoire à Séville, sur une analyse du programme d'études et sur des examens similaires en langue étrangère. Contrairement aux méthodes traditionnelles qui consistent à recueillir des évidences basées sur le contenu du cours pour les examens, ce procédé est réalisé avant d'écrire les différents points et tâches à réaliser de l'examen.

**Mots-clés:** Validité du contenu, examen d'anglais, premier cycle d'éducation secondaire, examen standardisé.

## 1. INTRODUCCIÓN

El grupo de investigación EALIS está desarrollando la prueba sumativa estandarizada

EALIS-ESO-2, cuya finalidad es evaluar, de acuerdo con lo establecido en el R.D. 1631/2006 (art. 3, España, 2007: 679), las capacidades de comprensión y expresión en la lengua inglesa de manera apropiada del alumnado de la ciudad de Sevilla al final del segundo curso de la etapa de la educación secundaria obligatoria (E.S.O.). En nuestra opinión, el contenido de esta prueba, es decir, «the themes, wording, and format of the items, tasks, or questions on a test, as well as the guidelines for procedures regarding administration and scoring» (American Educational Research Association, American Psychological Association y National Council on Measurement in Education, 1999: 11), debe tener en cuenta:

(i) Los objetivos y los contenidos del área de lengua extranjera para el segundo curso de la E.S.O., que en la Comunidad Autónoma de Andalucía están regulados por el Real Decreto 1631/2006 (España, 2007) y por la Orden de 10 de agosto de 2007 (Andalucía, 2007).

(ii) El contenido de las actividades de aprendizaje realizadas en el aula. Uno de los principios que debe regir la elaboración de una prueba sumativa es que la prueba debe incluir tareas que traten los mismos temas y que utilicen el mismo tipo de expresiones, instrucciones y formato que las actividades de aprendizaje realizadas en el aula. Si el contenido de una prueba sumativa consigue ajustarse al contenido de las actividades de aprendizaje realizadas en el aula, esto contribuirá, indudablemente, a la validación de la prueba (American Educational Research Association, American Psychological Association y National Council on Measurement in Education, 1999: 184).

(iii) El hecho de que la investigación en evaluación lingüística parece que ha demostrado que la actuación de los examinandos puede estar influida por el tipo de ítem o tarea que tiene que realizar. Shohamy (1983: 532-533), por ejemplo, sugería que las variables «tipo de actividad lingüística comunicativa» y «tema de la actividad» influían en las puntuaciones asignadas a los participantes en un experimento para evaluar sus competencias de expresión e interacción orales (*cf.* también, por ejemplo, Bachman y Palmer, 1982: 461-462). Este hecho viene a apoyar el principio de que la prueba sumativa debe incluir tareas similares a las actividades de aprendizaje realizadas en el aula.

En los estudios acerca de los métodos para obtener evidencias basadas en el contenido de la prueba (por ejemplo, Crocker, Miller y Franks, 1989; Osterlind, 1989; Oltman, Stricker y Barrows, 1990; Shavelson, Gao y Baxter, 1995; Sireci, 1998), estas evidencias se suelen obtener a partir del análisis de los juicios de una serie de jueces o mediante análisis estadísticos *después* de la redacción de los ítems o tareas. Así, Hambleton (1980) desarrolló un índice de congruencia ítem-objetivo (*item-objective congruence index*) para los ítems de pruebas criterioles que evalúan un único objetivo, cuyo valor dependía de la valoración realizada por una serie de especialistas en la materia del grado en que cada ítem medía cada objetivo. En este artículo se presenta un procedimiento para ajustar parte del contenido de las especificaciones de la prueba EALIS-ESO-2 a los objetivos y contenidos curriculares, es decir, *antes* de la redacción de los ítems y las tareas. Este

procedimiento, que fue ideado por Cuadrado Moreno (2009), consta de dos fases: (i) estudio observacional de clases de lengua inglesa del segundo curso de E.S.O. en la ciudad de Sevilla y (ii) elaboración de la especificación de la prueba. Con este procedimiento se pretende, en última instancia, solucionar un problema sobre el cual Popham (2000: 98) llamó la atención: la escasa calidad de las evidencias basadas en el contenido de la prueba que se proporcionan durante el proceso de validación de la prueba.

## **2. ESTUDIO OBSERVACIONAL DE CLASES DE LENGUA INGLESA DEL SEGUNDO CURSO DE E.S.O. EN LA CIUDAD DE SEVILLA**

El estudio observacional estuvo guiado por la siguiente pregunta de investigación:

• ¿Qué actividades de aprendizaje realizan durante la clase de lengua inglesa los alumnos sevillanos del segundo curso de E.S.O.?

El estudio observacional constó de las siguientes fases (Brown, 2001: 8-10):

- (i) Planificación del proyecto de observación.
- (ii) Diseño de un instrumento para la observación.
- (iii) Recogida y compilación de los datos de la observación.
- (iv) Análisis estadístico de los datos.

### **2.1. Planificación del proyecto de observación**

La conducta que debía ser observada era la actividad de aprendizaje que el alumnado sevillano del segundo curso de E.S.O. realiza durante la clase de lengua inglesa. El estudio observacional comenzó con una revisión de la literatura sobre la observación en el aula de segundas lenguas (Jarvis, 1968: 335-338; Clark, 1969: 392; Munby, 1978: 154-189; Johnson y Morrow, 1981; Malamah-Thomas, 1987; Allwright, 1988; Chaudron, 1988; Lier, 1988; Nunan, 1989; Allwright y Bailey, 1991; Wragg, 1994; Spada and Frohlich, 1995; Nikolov, 1999; Seedhouse, 2004).

La cuestión de cuántas categorías de la conducta observada debían distinguirse condujo al problema de determinar cuántas subpruebas debían componer la prueba EALIS-ESO-2. Partiendo de la declaración explícita del legislador español, según la cual éste había tomado como base para la elaboración del currículo del área de lengua extranjera en la E.S.O. el *Marco Común Europeo de Referencia para las Lenguas: Aprendizaje, Enseñanza, Evaluación* (M.C.E.R.) «Estas pautas [del M.C.E.R.] han sido un referente clave en este currículo [del área de lengua extranjera para la etapa de la E.S.O.]», (España, 2007: 742), se decidió que las subpruebas debían tener en cuenta las categorías del M.C.E.R. El problema que se debía solucionar a continuación era elegir como base de la prueba o el constructo de competencia lingüística comunicativa del M.C.E.R. (Trim,

North y Coste, 2002: 106-127) o el constructo de actividad lingüística comunicativa del M.C.E.R. (Trim, North y Coste, 2002: 60-88). Según Trim, North y Coste (2002), las competencias comunicativas de la lengua están compuestas de (i) las competencias lingüísticas, (ii) las competencias sociolingüísticas y (iii) las competencias pragmáticas (p. 106), mientras que para estos mismos autores existen los siguientes tipos de actividades lingüísticas comunicativas: (i) actividades de expresión, (ii) actividades receptivas, (iii) actividades interactivas, (iv) actividades de mediación, y (v) comunicación no verbal (pp. 60-88). En la traducción española del M.C.E.R., por cierto, se tradujeron los términos ingleses *production activities* y *receptive activities* (Trim, North and Coste, 2001: 57 y 65, respectivamente) como *actividades de expresión* y *actividades de comprensión*, respectivamente (Trim, North y Coste, 2002: 61 y 68, respectivamente). Puesto que en la redacción del objetivo *i* de la etapa de la E.S.O. «Comprender y expresarse en una o más lenguas extranjeras de manera apropiada», (España, 2007: 679), que está relacionado con la enseñanza/aprendizaje de las lenguas extranjeras, aparecen los mismos términos que aparecen en el M.C.E.R. para referirse a tipos de actividades lingüísticas comunicativas, se concluyó que el legislador español había decidido basar la organización del currículo en la taxonomía de actividades lingüísticas comunicativas del M.C.E.R. y, por tanto, se decidió usar esta taxonomía como criterio para la organización de la prueba EALIS-ESO-2.

Se acordó, asimismo, utilizar como modelo para la prueba EALIS-ESO-2 el examen *Flyers* (University of Cambridge ESOL Examinations, 2007: 16-26), puesto que *Flyers* (i) está dirigido a un grupo de edad similar (edad media 11.9 años, Marshall y Gutteridge, 2002: 5) y (ii) está compuesto de tres subpruebas (*Listening*, *Speaking* y *Reading & Writing*), que, en realidad, son lo que en la terminología del M.C.E.R. se denominan actividades de comprensión auditiva, comprensión lectora, producción e interacción orales y producción e interacción escritas. Por esta razón se decidió observar tres tipos diferentes de actividades de aprendizaje en el aula: (i) actividades de comprensión auditiva, (ii) actividades de comprensión de lectura y de expresión e interacción escritas y (iii) actividades de expresión e interacción orales.

## 2.2. Diseño de un instrumento para la observación

El instrumento utilizado para registrar las observaciones está basado en el modelo de las facetas de los métodos de una prueba (Bachean, 1990: 116-152) y en su desarrollo posterior, el modelo para la descripción de las tareas de una prueba de Bachman y Palmer (1996: 47-58). En la didáctica de segundas lenguas y en la evaluación lingüística se han elaborado diferentes instrumentos que se pueden utilizar para observar las actividades de aprendizaje en un aula de enseñanza de segundas lenguas y para describir las tareas de una prueba (Jarvis, 1968: 335-338; Clark, 1969: 392; Munby, 1978: 154-189; Allwright,

1988; Brown, 1989: 51-52; Spada and Frohlich, 1995; Norris, Brown, Hudson y Yoshioka, 1998: 69-97; Freedle y Kostin, 1999: 26-31; Nikolov, 1999; Weigle, 2002: 60-64; Carr, 2006; North, Figueras, Takala, Van Avermaet y Verhelst, 2009: 25-34). Sin embargo, hasta donde hemos podido saber, ninguno de ellos ha utilizado para la observación un número tan elevado de variables como el modelo de Bachman y Palmer (1996), que, además, puede ser utilizado para la descripción tanto de actividades de aprendizaje como de tareas en una prueba. Este modelo ha sido utilizado en diferentes estudios de evaluación lingüística, como Bachman, Davidson, Ryan y Choi (1995) o Bachman, Davidson y Milanovic (1996).

### **2.3 Recogida y compilación de los datos de la observación**

Los observadores coordinaron los procedimientos para la observación antes de realizar la observación en una serie de reuniones en las que se analizaron el modelo para la descripción de las tareas de una prueba de Bachman y Palmer (1996: 47-58) y el instrumento para la observación. A continuación, se observaron a lo largo del primer trimestre del curso 2007-2008 las actividades de aprendizaje realizadas en la clase de lengua inglesa (que suele tener una duración aproximada de una hora) en una muestra aleatoria simple proporcional estratificada de diecinueve grupos, seleccionados entre los trescientos grupos del segundo curso de E.S.O. que, según la Delegación Provincial de la Consejería de Educación en Sevilla, existían en la ciudad de Sevilla durante el primer trimestre del curso académico 2007-2008 (Kish, 1965: 75-86). Los estratos utilizados fueron: (i) grupos no bilingües de centros públicos, (ii) grupos bilingües de centros públicos y (iii) grupos de centros privados. Por tanto, la forma de la conducta observada fue no secuencial (Kitsantas, Ware y Kitsantas, 2005: 913-914).

### **2.4. Análisis estadístico de los datos recogidos y presentación de los datos**

Los datos recogidos durante el estudio observacional han sido publicados en Cuadrado Moreno, Reyes Fernández, Fernández Amaya, Sánchez Fernández, Fielden, Ocaña Escolar y Hernández López (2009), Fernández Amaya, Reyes Fernández, Fielden, Ocaña Escolar, Hernández López y Sánchez Fernández (2009) y Sánchez Fernández, Cuadrado Moreno, Reyes Fernández, Fernández Amaya, Fielden, Ocaña Escolar y Hernández López (2009).

## **3. ELABORACIÓN DE LAS TABLAS DE ESPECIFICACIONES DE LAS SUBPRUEBAS DE COMPRENSIÓN AUDITIVA Y DE EXPRESIÓN E INTERACCIÓN ORALES**

La elaboración de las tablas de especificaciones de la subpruebas estuvo guiada por

el procedimiento de Linn y Gronlund (2000: 143), que consta de tres fases: (i) elaboración de una lista de los objetivos de la enseñanza, (ii) elaboración de una lista con los contenidos del curso y (iii) elaboración de la tabla de las especificaciones.

Aunque durante el estudio observacional se habían utilizado tres categorías de actividades de aprendizaje, la revisión de la literatura acerca de las pruebas de aprovechamiento similares destinadas a un grupo de edad semejante y basadas en el M.C.E.R. (por ejemplo, France, 2007; Rupp, Vock, Harsch y Köller, 2007; Tuokko, 2007; Gassner, Mewald y Sigott, 2008; Gassner, Mewald y Sigott, s.d.; Mewald, Gassner y Sigott., s.d.a; Mewald, Gassner y Sigott, s.d.b) condujo a revisar la decisión de diseñar una prueba con tres subpruebas. Finalmente, se acordó que la prueba constaría de las siguientes subpruebas: (i) subprueba de comprensión escrita, (ii) subprueba de expresión e interacción escritas, (iii) subprueba de comprensión auditiva, y (iv) subprueba de expresión e interacción orales.

### **3.1. Elaboración de una lista de los objetivos de la enseñanza**

Los objetivos relacionados con las competencias lingüísticas comunicativas para la materia de lengua extranjera en el segundo curso de la E.S.O. aparecen en el Anexo II del Real Decreto 1631/2006, de 29 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la Educación Secundaria Obligatoria (España, 2006: 743), que establece, entre otras cosas, que la enseñanza de la lengua extranjera en la etapa de la educación secundaria obligatoria tendrá como finalidad el desarrollo de las siguientes capacidades [...]

1. Escuchar y comprender información general y específica de textos orales en situaciones comunicativas variadas [...]

2. Expresarse e interactuar oralmente en situaciones habituales de comunicación de forma comprensible, adecuada y con cierto nivel de autonomía [...]

Estos objetivos son desarrollados para el segundo curso de E.S.O. en la sección «Criterios de evaluación» (España, 2007: 746):

1. Comprender la idea general e informaciones específicas de textos orales emitidos por un interlocutor, o procedentes de distintos medios de comunicación, sobre temas conocidos [...]

2. Participar con progresiva autonomía en conversaciones y simulaciones relativas a las experiencias personales, planes y proyectos, empleando estructuras sencillas, las expresiones más usuales de relación social, y una pronunciación adecuada para lograr la comunicación [...]

Desafortunadamente, el R.D. 1631/2006 no establece un objetivo de nivel de competencia lingüística que el alumno debería alcanzar al final de la etapa de la E.S.O.

A partir de la revisión de otras pruebas dirigidas a alumnos del mismo grupo de edad (por ejemplo, France, 2007; Rupp, Vock, Harsch and Köller, 2007: 86-96; University of Cambridge ESOL, 2007: 3; Gassner, Mewald y Sigott, 2008: 20; Gassner, Mewald y Sigott, s.d.: 1; Mewald,, Gassner, y Sigott., s.d.a: 1; Mewald, Gassner and Sigott, s.d.b: 1), los autores decidieron establecer como objetivo para el alumnado al final de la etapa de E.S.O. un nivel de competencia lingüística comunicativa en la lengua extranjera correspondiente al nivel de referencia común A2 (Trim, North and Coste, 2001: 21-42). Esto conlleva que (i) el alumnado del segundo curso de E.S.O. debería tener una competencia lingüística en la lengua extranjera situada entre los niveles comunes de referencia A1 y A2 y (ii) la prueba de aprovechamiento EALIS-ESO-2 debe tener como objetivo determinar el grado de competencia lingüística comunicativa de los alumnos del segundo curso de E.S.O., que se debe encontrar entre los niveles comunes de referencia A1 y A2.

### **3.2. Elaboración de una lista con los contenidos del curso**

En el R.D. 1631/2006 los contenidos de la materia de lengua extranjera para el segundo curso de E.S.O aparecen distribuidos en «bloques», entre los que se encuentran el «Bloque 1. Escuchar, hablar y conversar» (España, 2007: 745):

Bloque 1. Escuchar, hablar y conversar

Escucha y comprensión de mensajes emitidos dentro del aula relacionados con las actividades habituales.

Obtención de información general y específica de textos orales sobre asuntos cotidianos y predecibles procedentes de diferentes medios de comunicación y con apoyo de elementos verbales y no verbales [...]

Producción de textos orales breves y coherentes sobre temas de interés personal y con pronunciación adecuada.

Participación en conversaciones y simulaciones dentro del aula, con pronunciación y entonación adecuadas para lograr la comunicación.

Empleo de respuestas adecuadas en situaciones de comunicación en el aula.

Desarrollo de estrategias de comunicación para superar las interrupciones en la comunicación y para iniciar y concluir intercambios comunicativos.

### **3.3. Elaboración de la tabla de las especificaciones**

El primer paso en la elaboración de las estructuras de las tablas de las especificaciones para las subpruebas consistió en establecer los objetivos de cada subprueba. Esta tarea fue realizada teniendo en cuenta (i) la distinción de Gronlund (2004: 13) entre objetivos

generales de la enseñanza y resultados específicos del aprendizaje, (ii) los contenidos que el R.D. 1631/2006 asigna al área de lengua extranjera en el segundo curso de E.S.O., y (iii) que el alumno debía alcanzar al final del segundo curso de E.S.O. un nivel común de referencia entre los niveles A1 y A2.

Gronlund (2004: 13) presentó una serie de definiciones relacionadas con los objetivos de la enseñanza:

1. *Objetivo general de la enseñanza (general instructional objective)*: «an intended outcome of instruction that has been stated in general enough terms to encompass a domain of student performance (e.g., «Comprehends the literal meaning of written material»).

2. *Resultado específico del aprendizaje (specific learning outcome)*: «an intended outcome of instruction that has been stated in terms of specific and observable student performance (e.g., «Identifies details that are explicitly stated in the passage»).

3. *Actuación del estudiante (student performance)*: «any measurable or observable student response that is the result of learning».

Los contenidos relacionados con las actividades de comprensión auditiva aparecen en el «Bloque 1 Escuchar, hablar y conversar» (España, 2007: 745):

Escucha y comprensión de mensajes emitidos dentro del aula relacionados con las actividades habituales.

Obtención de información general y específica de textos orales sobre asuntos cotidianos y predecibles procedentes de diferentes medios de comunicación y con apoyo de elementos verbales y no verbales.

Uso de estrategias de comprensión de los mensajes orales: uso del contexto verbal y no verbal y de los conocimientos previos sobre la situación, identificación de palabras clave.

Para la prueba EALIS-ESO-2-*Listening* decidimos centrarnos en el segundo de estos contenidos, que está relacionado con el criterio de evaluación 1 que el R.D. 1631/2006 establece para el segundo curso de E.S.O. «Comprender la idea general e informaciones específicas de textos orales emitidos por un interlocutor, o procedentes de distintos medios de comunicación, sobre temas conocidos», (España, 2007: 746). Como se puede apreciar, para el legislador español el concepto de comprensión auditiva incluye, por lo menos, dos subcompetencias diferentes: (i) la comprensión de la idea general en textos orales y (ii) la comprensión de informaciones específicas en textos orales. Por tanto el legislador ha adoptado una visión multidivisible de la comprensión auditiva, que tampoco está respaldada por la investigación empírica. De hecho, en la investigación sobre la comprensión auditiva no existe acuerdo acerca del número de subdestrezas relacionadas con la comprensión auditiva y, además, la mayor parte de las taxonomías de las subdestrezas relacionadas con la comprensión auditiva, como las de Munby (1978: 126-131) o Richards (1983: 228-230), carecen de una validación empírica. Resulta lamentable que el concepto de la comprensión auditiva adoptado por el legislador no tenga en

cuenta los resultados de investigaciones empíricas recientes, como las de Freedle and Kostin (1996), Nissan, DeVincenzi and Tang (1996) o Buck y Tatsuoka (1998). Por tanto, teniendo en cuenta la obligación de ceñirse a lo estipulado en el currículo de la materia de lengua extranjera, se decidió que la redacción de los resultados específicos del aprendizaje para la subprueba de comprensión auditiva quedara como sigue:

1. El estudiante identifica la información general en un texto oral.
2. El estudiante identifica la información específica en un texto oral.

Por último, para establecer los objetivos específicos de la subprueba de expresión e interacción orales se tuvo en cuenta:

(i) El objetivo general de la E.S.O. respecto a la materia de lengua extranjera «1. Comprender y expresarse en una o más lenguas extranjeras de manera apropiada», (España 2007: 679).

(ii) El objetivo 2 de la enseñanza de la materia de lengua extranjera «Expresarse e interactuar oralmente en situaciones habituales de comunicación de forma comprensible, adecuada y con cierto nivel de autonomía», (España, 2007: 743).

(iii) La redacción de los contenidos del «Bloque 1. Escuchar, hablar y conversar» para el segundo curso de la E.S.O. relacionados con las actividades lingüísticas comunicativas de la expresión e interacción orales (España, 2007: 745):

Producción de textos orales breves y coherentes sobre temas de interés personal y con pronunciación adecuada.

Participación en conversaciones y simulaciones dentro del aula, con pronunciación y entonación adecuadas para lograr la comunicación.

Empleo de respuestas adecuadas en situaciones de comunicación en el aula.

Desarrollo de estrategias de comunicación para superar las interrupciones en la comunicación y para iniciar y concluir intercambios comunicativos.

(iv) El criterio de evaluación 2 para el segundo curso de E.S.O. «Participar con progresiva autonomía en conversaciones y simulaciones relativas a las experiencias personales, planes y proyectos, empleando estructuras sencillas, las expresiones más usuales de relación social, y una pronunciación adecuada para lograr la comunicación» (España, 2007: 746).

Parece que el legislador español tuvo en cuenta la división del M.C.E.R. en actividades lingüísticas comunicativas de expresión e interacción orales (Trim, North and Coste, 2001: 58-60 y 73-81, respectivamente) puesto que (i) el contenido cuarto del Bloque 1 «Producción de textos orales breves y coherentes sobre temas de interés personal y con pronunciación adecuada» incluye el término *producción*, muy similar al término inglés *productive activity* que utiliza la edición del M.C.E.R. en inglés (Trim, North and Coste, 2001: 58) y (ii) el contenido quinto del Bloque 1 «Participación en conversaciones y simulaciones dentro del aula, con pronunciación y entonación adecuadas para lograr la

comunicación» incluye el término *conversación*, que es uno de los ejemplos de las actividades de interacción citados en el M.C.E.R. y el título de una de las escalas ilustrativas para valorar las actividades de interacción oral (Trim, North y Coste, 2002: 75 y 76-77, respectivamente). En la traducción del M.C.E.R. al español se tradujo el término inglés *oral production activities* (Trim, North and Coste, 2001: 58) como *actividades de expresión oral* (Trim, North y Coste, 2002: 61), es decir, parece que a la hora de redactar el contenido «Producción de textos orales breves y coherentes sobre temas de interés personal y con pronunciación adecuada» el legislador español utilizó el término *producción* influido por el término inglés original *production*. Puesto que en el M.C.E.R. se distingue entre actividades de expresión oral y actividades de interacción orales, y la participación en conversaciones está claramente relacionada con las actividades de interacción oral (Trim, North and Coste, 2001: 73), concluimos que el contenido de la producción de textos orales breves y coherentes estaba relacionado con las actividades de expresión oral. Además, el criterio de evaluación 2 obliga claramente a que los alumnos realicen actividades de actuación. Asimismo, los resultados de la observación de las actividades de aprendizaje en el aula mostraron que los alumnos del segundo curso de E.S.O. realizaban tanto actividades de expresión oral como actividades de interacción oral, aunque la inmensa mayoría de éstas eran actividades de interacción oral. Por tanto, se acordó establecer los siguientes resultados específicos del aprendizaje para la subprueba de expresión e interacción orales:

1. El alumno produce textos orales breves.
2. El alumno participa en conversaciones.

### 3.3.1. Duración de las subpruebas

En primer lugar, aparecen una serie de tablas que presentan la duración de subpruebas similares a las que estamos diseñando:

Prueba	Duración
<i>DELFA-2 Version Scolaire et Junior - Compréhension de l'oral</i> (Centre International d'Études Pédagogiques s.d.: 1)	25
<i>Flyers - Listening</i> (University of Cambridge ESOL Examinations 2007: 27)	25
<i>E8-Standards Listening Test</i> (Mewald, Gassner and Sigott s.d.a: 4)	35

Tabla 1 *Duración de diversas subpruebas similares de comprensión auditiva.*

Prueba	Duración
<i>DELFA-2 Version Scolaire et Junior – Production orale</i> (Centre International d'Études Pédagogiques s.d.: 1)	8
<i>Flyers – Speaking</i> (University of Cambridge ESOL Examinations 2007: 31)	7-9
<i>E8-Standards Speaking Test</i> (realizada simultáneamente por dos alumnos) (Mewald, Gassner and Sigott s.d.b: 4)	14

Tabla 2 *Duración de diversas subpruebas similares de expresión e interacción orales.*

La duración de una prueba está muy determinada, entre otras cosas, por el tiempo disponible para la convocatoria (Schmeiser and Welch, 2006: 311) y el grado necesario de fiabilidad (Hambleton, Swaminathan, Algina y Coulson, 1978: 23-24), que, a su vez, está influido por el grado de importancia de la decisión que se toma a partir de la puntuación en la prueba (Linn y Gronlund, 2000: 130). Si se acepta que el grado de importancia de la decisión es elevado y que, como máximo, se dispondría de dos horas para realizar las subpruebas de comprensión escrita, producción e interacción escritas y comprensión auditiva, esto obliga a repartir dos horas entre estas subpruebas. Las características prácticas de los centros que imparten E.S.O., con clases con una duración aproximada de una hora, hacen aconsejable no dedicar más de 30-40 minutos a cualquier subprueba. A partir de las modas de los valores de la variable «Duración de una prueba», se acordó asignar las siguientes duraciones a las subpruebas de comprensión auditiva y de expresión e interacción orales:

Subprueba	Duración
Comprensión auditiva	25
Expresión e interacción orales	7

Tabla 3 *Duración de las subpruebas de comprensión auditiva y de expresión e interacción orales EALIS-ESO-2.*

### *3.3.2. Estructura de las subpruebas*

Para cada subprueba se presenta, en primer lugar, una tabla con la estructura de subpruebas similares y, a continuación, se justifica y se aporta la propuesta de especificación para esa subprueba.

La Tabla 4 expone la estructura de subpruebas similares de comprensión auditiva:

Subprueba	Número de tareas	Número total de ítems	Media de ítems por tarea
<i>DELFA-2 Version Scolaire et Junior – Compréhension de l’oral</i> (Centre International d’Études Pédagogiques s.d.: 1)	3	14	4
E8-Standards Listening Test (Mewald, Gassner and Sigott s.d.a: 4)	3	20	6

Tabla 4 *Estructura de diversas subpruebas similares de comprensión auditiva.*

Basándose en el hecho de que en las actividades observadas de comprensión auditiva la moda en la variable *número de ítems en cada actividad de aprendizaje* era 5, se llegó al acuerdo de asignar cinco ítems a cada tarea de la subprueba y distribuir los veinticinco minutos que debía durar esta subprueba de un modo uniforme, es decir, cada tarea debía durar aproximadamente ocho minutos. Esto llevó a la siguiente estructura de la tabla de especificaciones para la subprueba EALIS-ESO-2-*Listening* (Tabla 5):

Tarea	Subprueba	Número de ítems	Duración
1	El estudiante identifica la información general en un texto oral El estudiante identifica la información específica en un texto oral	1 4	8
2	El estudiante identifica la información general en un texto oral El estudiante identifica la información específica en un texto oral	1 4	8
3	El estudiante identifica la información general en un texto oral El estudiante identifica la información específica en un texto oral	1 4	8

Tabla 5. *Estructura de la subprueba de comprensión auditiva EALIS-ESO-2-Listening.*

La Tabla 6 muestra la estructura de subpruebas similares de expresión e interacción orales:

Subprueba	Número de tareas
<i>DELFA-2 Version Scolaire et Junior – Production orale</i> (Centre International d’Études Pédagogiques s.d.: 1)	3
<i>Flyers – Speaking</i> (University of Cambridge ESOL Examinations 2007: 31)	4
<i>E8-Standards Speaking Test</i> (Mewald, Gassner and Sigott s.d.b: 5)	3

Tabla 6 *Estructura de diversas subpruebas similares de expresión e interacción orales.*

Como se aprecia en la Tabla 6, la moda en la variable *número de tareas* es 3, lo cual condujo a proponer una subprueba de expresión e interacción orales con tres actividades de expresión e interacción orales, cuya estructura quedó como sigue (Tabla 7):

Tarea	Objetivo	Duración
1	El alumno participa en conversaciones	8
2	El alumno produce textos orales breves	8
3	El alumno participa en conversaciones	8

Tabla 7. *Estructura de la subprueba EALIS-ESO-2-Speaking.*

#### 4. CONCLUSIÓN

Este estudio ha presentado un procedimiento para el desarrollo de una parte de la especificación de las subpruebas de comprensión auditiva y de expresión e interacción orales de una prueba de aprovechamiento de lengua inglesa (EALIS-ESO-2) para el alumnado del segundo curso de E.S.O. de la ciudad de Sevilla. El diseño está basado en (i) un estudio observacional de las actividades de aprendizaje que el alumnado del segundo curso de E.S.O. de la ciudad de Sevilla realizaba durante la clase de lengua inglesa y (ii) un examen del currículo de la materia de lengua extranjera para el segundo curso de E.S.O. (España 2007) y de los presupuestos que subyacen a este currículo, lo cual constituye una alternativa a los métodos tradicionales para obtener evidencias basadas en el contenido de la prueba y una solución al problema de la escasa calidad de este tipo de evidencias.

#### REFERENCIAS BIBLIOGRÁFICAS

- ALLWRIGHT, Dick. (1988): *Observation in the Language Classroom*. London y New York, Longman.
- ALLWRIGHT, Dick. y BAILEY, Kathleen M. (1991): *Focus on the Language Classroom: An Introduction to Classroom Research for Language Teachers*. Cambridge, Cambridge University Press.
- AMERICAN EDUCATIONAL RESEARCH ASSOCIATION, AMERICAN PSYCHOLOGICAL ASSOCIATION y NATIONAL COUNCIL ON MEASUREMENT IN EDUCATION (1999): *Standards for Educational and Psychological Testing*. Washington, D.C., American Educational Research Association.
- ANDALUCÍA. (2007): Orden de 10 de agosto de 2007, por la que se desarrolla el

- currículo correspondiente a la Educación Secundaria Obligatoria en Andalucía. *Boletín Oficial de la Junta de Andalucía* 171, pp. 23-65.
- BACHMAN, Lyle F. (1990): *Fundamental Considerations in Language Testing*. Oxford, Oxford University Press.
- BACHMAN, Lyle F., DAVIDSON, Fred y MILANOVIC, Michael (1996): «The use of test method characteristics in the content analysis and design of EFL proficiency tests», en *Language Testing* 13, pp. 125-150.
- BACHMAN, Lyle F., DAVIDSON, Fred, RYAN, Katherine y CHOI, Inn-Chull (1995): *An Investigation into the Comparability of Two Tests of English as a Foreign Language*. Cambridge, Cambridge University Press.
- BACHMAN, Lyle F. y PALMER, Adrian S. (1982): «The construct validation of some components of communicative proficiency», en *TESOL Quarterly* 16, pp. 449-465.
- BACHMAN, Lyle F. y PALMER, Adrian S. (1996): *Language Testing in Practice*. Oxford, Oxford University Press.
- BROWN, James Dean (1989): «Cloze item difficulty», en *JALT Journal* 11, pp. 46-67.
- BROWN, James Dean (2001): *Using Surveys in Language Programs*. Cambridge, Cambridge University Press.
- BUCK, Gary (2001): *Assessing Listening*. Cambridge, Cambridge University Press.
- BUCK, Gary y TATSUOKA, Kikumi (1998): «Application of the rule-space procedure to language testing: examining attributes of a free response listening test», en *Language Testing* 15, pp. 119-157.
- CARR, Nathan T. (2006): «The factor structure of test task characteristics and examinee performance», en *Language Testing* 23, pp. 269-289.
- CENTRE INTERNATIONAL D'ÉTUDES PEDAGOGIQUES (s.d.): *Diplôme d'Études en Langue Française DELF A-2: Version Scolaire et Junior*. Paris, Centre International d'Études Pédagogiques.
- CHAUDRON, Craig (1988): *Second Language Classrooms: Research on Teaching and Learning*. Cambridge, Cambridge University Press.
- CLARK, John L. D. (1969): «The Pennsylvania Project and the 'audio-lingual vs. traditional' question», en *The Modern Language Journal* 53, pp. 388-396.
- CROCKER, Linda M., MILLER, M. David y FRANKS, Elizabeth Ann (1989): «Quantitative methods for assessing the fit between test and curriculum», en *Applied Measurement in Education* 2, pp. 179-194.
- CUADRADO MORENO, José (2009): «Evidencias basadas en el contenido para las pruebas estandarizadas en primeras y segundas lenguas en entornos escolares formales», en BRETONES, Carmen et al. (eds.): *Applied Linguistics Now: Understanding Language and Mind/ La Lingüística Aplicada Actual: comprendiendo el Lenguaje y la Mente*. Almería, Universidad de Almería.

- ESPAÑA (2007): Real Decreto 1631/2006 de 29 de diciembre, por el que se establecen las enseñanzas mínimas correspondientes a la educación secundaria obligatoria. *Boletín Oficial del Estado* 5, pp. 677-773.
- FERNÁNDEZ AMAYA, Lucía, REYES FERNÁNDEZ, María, FIELDEN, Laura, OCAÑA ESCOLAR, Pablo, HERNÁNDEZ LÓPEZ, María de la O, y SÁNCHEZ FERNÁNDEZ, Beatriz (2009): «La validación del contenido de las subpruebas orales de la prueba de aprovechamiento EALIS-ESO-2 para la evaluación de la competencia en lengua inglesa del alumnado del segundo curso de la educación secundaria obligatoria», en BRETONES, Carmen et al. (eds.): *Applied Linguistics Now: Understanding Language and Mind/ La Lingüística Aplicada Actual: comprendiendo el Lenguaje y la Mente*. Almería, Universidad de Almería.
- FITZPATRICK, Robert y MORRISON, Edward J. (1971): «Performance and product evaluation», en THORNDIKE, Robert L. (ed.): *Educational Measurement*. Washington, D.C., American Council on Education, pp. 237-270.
- FRANCE (2007): Arrêté du 15-5-2007: Modalités d'attribution du diplôme national du brevet. *Bulletin Officielle*, 22. Disponible en Internet (15.5.2007): <http://www.education.gouv.fr/bo/2007/22/MENE0753209A.htm>.
- FREEDLE, Roy y KOSTIN, Irene (1996): *The Prediction of TOEFL Listening Comprehension Item Difficulty for Mini-talk Passages: Implications for Construct Validity*. Princeton, New Jersey, Educational Testing Service.
- FREEDLE, Roy y KOSTIN, Irene (1999): «Does the text matter in a multiple-choice test of comprehension? The case for the construct validity of TOEFL's minitalks», en *Language Testing* 16, pp. 2-32.
- GOUGH, Philip B., JUEL, Connie y GRIFFITH, Priscilla L. (1992): «Reading, speaking and the orthographic cipher», en GOUGH, P., EHRI, L. y TREIMAN, R. (eds.): *Hillsdale Reading Acquisition*. New Jersey, Lawrence Erlbaum, pp. 35-48.
- HAMBLETON, Ronald K. (1980): «Test score validity and standard-setting methods», en BERK, R. A. (ed.): *Criterion-referenced Measurement: The State of the Art*. Baltimore, John Hopkins University Press, pp. 80-123.
- HAMBLETON, Ronald K., SWAMINATHAN, Hariharan, ALGINA, James y COULSON, Douglas Bill (1978): «Criterion-referenced measurement and testing: a review of technical issues and developments», en *Review of Educational Research* 48, pp. 1-47.
- HAMBLETON, Ronald K. y ZENISKY, April (2003): «Advances in criterion-referenced testing methods and practices», en REYNOLDS, Cecil R. y KAMPHAUS, Randy W. (eds.): *Handbook of Psychological and Educational Assessment of Children: Intelligence, Aptitude and Achievement*. New York, Guilford Press, pp. 377-404.
- JARVIS, Gilbert A. (1968): «A behavioral observation system for classroom foreign

- language skill acquisition activities», en *The Modern Language Journal* 52, pp. 335-341.
- JOHNSON, Keith y MORROW, Keith (1981): *Communication in the Classroom: Applications and Methods for a Communicative Approach*. London, Longman.
- KISH, Leslie (1965): *Survey Sampling*. New York, John Wiley and Sons.
- KITSANTAS, Anastasia, WARE, Herbert W. y KITSANTAS, Panagiota (2005): «Observational studies», en KEMPF-LEONARD, Kimberly (ed.): *Encyclopedia of Social Measurement: Vol. 2 F-O.*, Oxford, Elsevier, pp. 913-918.
- LIER, Leo van (1988): *The Classroom and the Language Learner: Ethnography and Second Language Classroom Research*. London, New York, Longman.
- LINN, Robert L. y GRONLUND, Norman E. (2000): *Measurement and Assessment in Teaching*. Saddle River, New Jersey, Prentice-Hall.
- MALAMAH-THOMAS, Ann. (1987): *Classroom Interaction*. Oxford, Oxford University Press.
- MARSHALL, Helen y GUTTERIDGE, Mike (2002): «Candidate performance in the Young Learner English Tests in 2000», en *Research Notes* 7, pp. 5-8.
- MEWALD, Claudia, GASSNER, Otmar y SIGOTTI, Guenther (s.d.): *E8 Listening Test Specifications Version 03*. Klagenfurt, Language Testing Centre, Alpen-Adria-Universität.
- MEWALD, Claudia, GASSNER, Otmar y SIGOTTI, Guenther (s.d.): *E8 Speaking Test Specifications Version 01*. Klagenfurt, Language Testing Centre, Alpen-Adria-Universität.
- MUNBY, John (1978): *Communicative Syllabus Design: A Sociolinguistic Model for Defining the Content of Purpose-specific Language Programmes*. Cambridge, Cambridge University Press.
- NIKOLOV, Marianne (1999): «Classroom observation project», en FEKETE, Hajnal, MAJOR, Eva y NIKOLOV, Marianne (eds.): *English Language Education in Hungary: A Baseline Study*. Budapest, The British Council, pp. 221-246.
- NISSAN, Susan, DEVINCENZI, Felicia y TANG, K. Linda (1996): *An Analysis of Factors Affecting the Difficulty of Dialogue Items in TOEFL Listening Comprehension*. Princeton, New Jersey, Educational Testing Service.
- NORRIS, John M., BROWN, James Dean, HUDSON, Thom D. y J. YOSHIOKA, Jim (1998): *Designing second language performance assessments*. Honolulu, Hawai'i, Second Language Teaching and Curriculum Center, University of Hawai'i Press.
- NORTH, Brian, FIGUERAS, Neus, TAKALA, Sauli, VAN AVERMAET, Piet y VERHELST, Norman (2009): *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR): A Manual*. Strasbourg, Council of Europe, Language Policy Division.
- NUNAN, David (1989): *Understanding Language Classrooms: A Guide for Teacher-Initiated Action*. London, Prentice-Hall.
- OLTMAN, Philip K., STRICKER, Lawrence J. y BARROWS, Thomas S. (1990): «Analyzing test structure by multidimensional scaling», en *Journal of Applied Psychology*

75, pp. 21–27.

- OSTERLIND, Steven J. (1989): *Constructing Test Items*. Hingham, Massachusetts, Kluwer.
- POPHAM, W. James (2000): *Modern Educational Measurement: Practical Guidelines for Educational Leaders*. Boston, Allyn and Bacon.
- RICHARDS, Jack C. (1983): «Listening comprehension: Approach, design, procedure», en *TESOL Quarterly* 17, pp. 219-240.
- ROST, Detlef H. (1993): «Assessing the different components of reading comprehension: fact or fantasy?», en *Language Testing* 10, pp. 79-92.
- RUPP, André A., VOCK, Miriam, HARSCH, Claudia y KÖLLER, Olaf (2007): *The Development, Calibration, and Validation of Standards-based Tests for English as a First Foreign Language at the IQB: Part I –The Context for and Processes of Task Development*. Berlin, Institut zur Qualitätsentwicklung im Bildungswesen.
- SÁNCHEZ FERNÁNDEZ, Beatriz, CUADRADO MORENO, José, REYES FERNÁNDEZ, María, FERNÁNDEZ AMAYA, Lucía, FIELDEN, Laura, OCAÑA ESCOLAR, Pablo y HERNÁNDEZ LÓPEZ, María de la O (2009): «La validación del contenido de la subprueba de recepción y producción escritas de la prueba de aprovechamiento EALIS-ESO-2 para la evaluación de la competencia en lengua inglesa del alumnado del segundo curso de la educación secundaria obligatoria», en BRETONES, Carmen et al. (eds.): *Applied Linguistics Now: Understanding Language and Mind/ La Lingüística Aplicada Actual: comprendiendo el Lenguaje y la Mente*. Almería, Universidad de Almería.
- SCHMEISER, Cynthia B. y WELCH, Catherine J. (2006), «Test development», en BRENNAN, R. L. (ed.): *Educational Measurement*. Westport, Connecticut, Praeger Publishers, pp. 307-353.
- SEEDHOUSE, Paul (2004): *The Interactional Architecture of the Language Classroom: A Conversation Analysis Perspective*. Malden, Massachusetts, Blackwell.
- SHAVELSON, Richard J., GAO, Xiaohong y BAXTER, Gail P. (1995): «On the content validity of performance assessments: Centrality of domain specification», en BIRENBAUM, Menucha y DOUCHY, Filip J. R. C. (eds.): *Alternatives in Assessment of Achievements, Learning Process and Prior Knowledge*. Boston, Kluwer Academic, pp. 131–141.
- SHOHAMY, Elana (1983): «The stability of oral proficiency assessment on the oral interview testing procedure», en *Language Learning* 33, pp. 527-540.
- SIRECI, Stephen G. (1998): «The construct of content validity», en *Social Indicators Research* 45, pp. 83-117.
- SPADA, Nina y FROHLICH, Maria (1995): *Communicative Orientation of Language Teaching Observation Scheme: Coding Conventions and Applications*. Sydney, National Centre for English Language Teaching and Research.

- TRIM, John L. M., NORTH, Brian y COSTE, Daniel (2001): *Common European Framework of Reference for Languages*. Strasbourg: Council of Europe. Disponible en Internet (25.7. 2007): [http://www.coe.int/T/DG4/Linguistic/Source/Framework\\_EN.pdf](http://www.coe.int/T/DG4/Linguistic/Source/Framework_EN.pdf).
- TRIM, John L. M., NORTH, Brian y COSTE, David (2002): *Marco Común Europeo de Referencia: Aprendizaje, Enseñanza, Evaluación*. Madrid, Secretaría General Técnica del Ministerio de Educación, Cultura y Deporte y Grupo Anaya. Disponible en Internet (3.7.2009): [http://cvc.cervantes.es/obref/marco/cvc\\_mer.pdf](http://cvc.cervantes.es/obref/marco/cvc_mer.pdf).
- TUOKKO, Eeva (2007) : Mille tasolle perusopetuksen englannin opiskelussa päästään? Perusopetuksen päättövaiheen kansallisen arvioinnin 1999 Eurooppalaisen viitekehyksen taitotasoihin linkitetyt tulokset [What Levels do pupils reach in English at the end of the Comprehensive School? National Assessment Results linked to the Common European Framework]. Jyväskylä, Jyväskylä, University Publishing House.
- UNIVERSITY OF CAMBRIDGE ESOL EXAMINATIONS (2007): *Cambridge Young Learners English Tests: Starters | Movers | Flyers: Handbook for Teachers*. Cambridge, University of Cambridge ESOL Examinations.
- WRAGG, Edward Conrad (1999): *An Introduction to Classroom Observation*. London, Routledge.