

## PRUEBAS DE EVALUACIÓN EN INGLÉS PARA FINES ESPECÍFICOS (IFE): TESTS DE MICRO-DESTREZAS LINGÜÍSTICAS/ TESTS DE MACRO-DESTREZAS COMUNICATIVAS

M<sup>a</sup> MERCEDES RICO GARCÍA  
*Universidad de Extremadura*

**RESUMEN.** *La aparición y desarrollo del Inglés para Fines Específicos (IFE -ESP: English for Specific Purposes en terminología inglesa-), viene determinada por el aumento producido en los últimos años de profesionales que necesitan dominar la lengua inglesa en áreas, contextos y campos específicos.*

*Sin embargo, la rápida expansión del inglés específico no vino acompañada de un desarrollo paralelo de fórmulas de evaluación en IFE, circunstancia que justifica, entre otras, la formulación teórica y la investigación empírica de lo que consideramos una parte integrante del proceso de enseñanza/aprendizaje de las lenguas específicas. El objetivo del presente artículo se centra en presentar una revisión teórica, adecuándola a nuestro contexto –inglés para fines específicos como idioma extranjero- de dos de los tipos de pruebas de evaluación más utilizadas: pruebas de evaluación de habilidades lingüísticas (micro-destrezas lingüísticas) y pruebas de evaluación de la competencia comunicativa (macro-capacidades comunicativas).*

**PALABRAS CLAVE.** *Inglés específico, pruebas de evaluación en IFE, tests de micro-capacidades, tests de macro-capacidades comunicativas*

**ABSTRACT.** *The expansion created in scientific, technical and economic activities generated a demand for an international language for specific purposes. This increasing awareness has created a new generation of professionals who need to learn English for definite reasons and specific circumstances.*

*However, the rapid expansion in ESP (English for Specific Purposes) was not accompanied by a similar increase in new ESP testing formulae, reason why theoretical studies and empirical research on language users is necessary. Being considered as a crucial and integral part of the teaching/learning process in ESP, the purpose of this article is focused on presenting a theoretical revision in our context -English for Specific Purposes as a Foreign Language- of two of the most usual tests: achievement tests and proficiency tests.*

**KEYWORDS.** *English for Specific Purposes, ESP evaluation, achievement tests, proficiency tests.*

## 1. EVALUACIÓN EN IFE: CONCEPTOS FUNDAMENTALES

La necesidad de aprender inglés por motivos específicos (Hutchinson y Waters 1987), coincidiendo con el desarrollo de los métodos comunicativos de aprendizaje, condujo prògresivamente a la elaboración de material específico adecuado a las diferentes necesidades comunicativas surgidas en campos científico-técnicos.

Así, junto con el estudio sistemático de la lengua, en la década de los 70 comenzó también a adquirir importancia el análisis de las necesidades comunicativas de grupos de personas que por razones académicas y/o profesionales necesitaban dominar la lengua inglesa en contextos específicos. El análisis de necesidades se convertiría así en uno de los rasgos más importantes de la fundamentación de IFE, como el propuesto por Candlin, Leather y Bruton (1976) basado en la identificación de funciones y nociones lingüísticas, o el más que celebrado modelo de Mumby (1978), en donde se nos presentaba un sistema detallado de correspondencias entre las necesidades de discentes específicos con el contenido de los programas de formación en lengua inglesa “*devising appropriate syllabus specifications from adequate profiles of communications needs*” (Mumby 1978: 3).

Consideramos que el principal propósito de la enseñanza/aprendizaje de las Lenguas para Fines Específicos se basa en la estrecha relación que debe establecerse entre la formación académica de los discentes y los usos comunicativos concretos, en la necesidad de adecuar los programas de enseñanza a las exigencias académico-profesionales de grupos específicos de usuarios.

La rápida expansión de cursos de inglés especializados no vino acompañada, sin embargo, de la formulación de nuevos métodos y modelos de evaluación, no siendo hasta la aparición del Servicio de Evaluación de la Lengua Inglesa (ELTS: *English Language Testing System*) en 1980 cuando comenzara a percibirse la necesidad de elaborar pruebas en lengua inglesa pertenecientes a las diferentes áreas de especialización y a establecerse, por tanto, los criterios fundamentales sobre los que se asienta la evaluación de lenguas<sup>1</sup>.

La falta de conceptualización teórica no ha impedido, sin embargo, que en los últimos años se hayan diseñado pruebas de evaluación específicas, destacando, por ejemplo, el listado que presentan Davies y West (1984) sobre 14 pruebas de evaluación en IFE administradas por diferentes instituciones británicas, como el test *English for Business and English For Science* de la Universidad de Cambridge, las llevadas a cabo por la Cámara de Comercio e Industria de Londres sobre inglés empresarial o un largo etcétera de pruebas diferentes a nivel mundial diseñadas para una gran variedad de especialidades diversas.

Así, diferentes propósitos de evaluación determinarán, sin duda, la elaboración de tipos de pruebas diferentes: pruebas de aptitud en ámbitos profesionales, pruebas de acceso a cursos de formación o pruebas a nivel académico elaboradas y administradas con propósitos específicos (diagnosticar carencias comunicativas, establecer el nivel de competencia lingüística y similares).

Thus we can speak of selection, entrance and readiness tests with regard to admission, decision, placement and diagnostic tests with regard to identifying the appropriate instructional level or specific areas in which construction is needed and progress, achievement, attainment, or mastery tests with respect to decisions about how individuals should proceed through the program, or how well they are attaining the program's objectives. (Bachman 1990: 70).

## 2. PRUEBAS DE EVALUACIÓN EN IFE

Aunque diseñadas con objetivos iniciales diferentes, existe una superposición en cuanto a la finalidad e información extraída a partir de los diferentes tipos de pruebas de evaluación. Así, y a pesar de la multifuncionalidad de cada una de ellas, las pruebas de evaluación podrían clasificarse atendiendo a cinco criterios principales<sup>2</sup>:

1. Según el uso, podrían dividirse en pruebas iniciales, pruebas de selección, de diagnóstico y de evaluación de la competencia lingüística.
2. Según el contenido, se clasificarían en pruebas de evaluación de habilidades lingüísticas específicas (tests de micro-destrezas lingüísticas) y pruebas globales de evaluación (tests de macro-destrezas comunicativas).
3. De acuerdo con el marco de referencia a la hora de interpretar los resultados, las clasificaríamos en pruebas de evaluación normalizadas (*Norm-referenced*) y pruebas de evaluación atendiendo a criterios de nivel comunicativo (*Criterion-referenced*)<sup>3</sup>.
4. En cuanto al método de puntuación, podemos distinguir entre métodos de puntuación objetivos (como es el caso, por ejemplo, de los tests de opción múltiple o los tests *cloze*), y métodos de puntuación subjetivos en los que, como ocurre en el caso de las redacciones, resúmenes, entrevistas orales, etc., no existe un procedimiento objetivo de puntuación, dependiendo, en ocasiones, de la opinión del profesor o tribunal la decisión de si la prueba ha sido realizada satisfactoriamente por los discentes.
5. La última característica en la que podrían basarse las diferencias entre los tipos de pruebas es la referente al test específico utilizado, y entre los que podríamos distinguir los tests de opción múltiple, resúmenes, tests *cloze*, entrevistas orales, etc.

Asimismo, con el fin de minimizar los efectos de las limitaciones de fiabilidad y validez subyacentes en las pruebas de evaluación, deberían vigilarse siempre tres puntos fundamentales en el diseño y administración de las mismas<sup>4</sup>:

- 1). Establecer con claridad los propósitos, consecuencias y proyecciones a partir de los resultados obtenidos -ya sea una prueba realizada en ámbitos académicos como en ámbitos profesionales-, y los criterios de admisión y selección de los participantes.
- 2). Definir las habilidades comunicativas a evaluar: identificación de los usos comunicativos en contextos específicos (*Tasks in the TLU: Target Language Use*).

Es necesario señalar que varias son las fuentes que deben emplearse en la identificación y selección de los usos y habilidades específicas en IFE. Así, destacaríamos (1) el análisis de necesidades de los discentes, estudio que aportaría una información esencial en la selección y elaboración de los contenidos de los cursos de inglés específicos (*needs, lacks and wants*), (2) la opinión del profesor de lenguas, opinión que, atendiendo a las necesidades lingüístico/comunicativas, dependería también de lo familiarizado que estuviera con el área de especialización de los discentes y (3) la información extraída a partir de las entrevistas con profesores especialistas del área y profesionales específicos a nivel laboral. En este contexto, no podemos olvidar las dos áreas principales sobre las que se proyecta la enseñanza del inglés específico: la enseñanza del inglés a nivel académico y el uso de la lengua en ámbitos profesionales.

3). Describir las condiciones y contenidos de las pruebas.

Las consideraciones referentes a la estructura y organización de las pruebas de evaluación: partes, relación entre las mismas, importancia de cada una de ellas, propósito, pautas de realización, duración y puntuación, etc. deben establecerse siempre de forma pormenorizada.

Coincidiendo con la consideración de Bachman (1990:71) en cuanto a la clasificación de las pruebas de evaluación "Language tests can be based on either a theory of language proficiency or a specific domain of content (...) while syllabus-based tests are generally referred to as achievement tests", consideramos conveniente realizar las pruebas de evaluación atendiendo a dos criterios principales: (1) a partir de la información recogida en la programación general, *syllabus-based-construct definitions*, pruebas adecuadas para obtener información acerca de la competencia de los alumnos sobre áreas lingüísticas concretas: pruebas iniciales y de diagnóstico diseñadas para alumnos que acceden, por ejemplo, a un programa de producción escrita a nivel universitario; test de micro-capacidades gramaticales enfocado a evaluar estructuras recogidas en la programación, conocimiento de vocabulario específico, organización de párrafos, etc.; (2) a partir de un modelo teórico para establecer el nivel de competencia específica, modelo no basado en los contenidos puntuales de la programación sino en la evaluación de la capacidad comunicativa misma, *language proficiency*: pruebas iniciales para licenciados en informática aspirantes, por ejemplo, a un puesto en una empresa de programación, tests de macro-destrezas comunicativas para evaluar los resultados de un curso de producción oral, test inicial diseñado para un curso de comprensión lectora en IFE y similares.

Presentamos a continuación algunas de las características principales de las pruebas de evaluación atendiendo al contenido de las mismas: pruebas de evaluación de micro-capacidades lingüísticas (*achievement tests*) y pruebas de macro-capacidades comunicativas (*proficiency tests*).

### 3. PRUEBAS DE EVALUACIÓN DE MICRO-CAPACIDADES LINGÜÍSTICAS

Coincidiendo con gran parte de los investigadores del área, consideramos, en primer lugar, que las pruebas de evaluación de habilidades específicas -evaluación de

micro-capacidades- se basarían en la evaluación de los contenidos programados para el desarrollo de los cursos.

This kind of test is the least problematic, since it is usually internal to the course. It does not have to conform to external influences, but should rather reflect the nature and content of the course itself. (Alderson y Hughes 1981).

These test how well the learner is keeping up with the syllabus and can be administered any time through the course. (Hutchinson y Waters 1987).

Sin establecer un número determinado, y realizándose tantas veces como fueran necesarias a lo largo del curso, las pruebas de evaluación de las micro-capacidades nos informarían, en primer lugar, del desarrollo y evolución de las clases, del nivel de competencia que los discentes van adquiriendo en los diferentes usos lingüísticos recogidos en el temario y, compartiendo funciones con los tests de diagnóstico, de las habilidades en las que se deberían seguir trabajando.

Sin embargo, y aunque en la propia definición de las mismas se recoge la posibilidad de evaluar habilidades lingüísticas aisladas: capacidad para formar interrogativas, deducir significado por el contexto, pronunciación de la terminación *ed* de los pasados y los participios de los verbos regulares, etc., advertimos de la inconveniencia de utilizar este tipo de parámetros o micro-destrezas como único criterio de evaluación. Encontramos con demasiada frecuencia que muchas de las pruebas finales de evaluación, en todos los niveles de enseñanza de la lengua, se elaboran, casi en su totalidad, a partir de elementos pertenecientes al nivel de las micro-capacidades, circunstancia que conlleva la falta de referencia en la evaluación de habilidades comunicativas por parte del profesor y la pérdida de la utilización de competencias comunicativas específicas por parte de los discentes. Introducir exclusivamente en un test de comprensión lectora, por ejemplo, preguntas referentes al vocabulario específico, correspondencia sintáctica o gramatical, relación existente entre cada una de las partes del texto, etc., nos conduciría a la pérdida de referencia del nivel en comprensión lectora de los discentes, imposibilitando la capacidad de evaluar el nivel de competencia real en dicha destreza. Opinamos que, aun interesados en la evaluación de las micro-capacidades lingüísticas, un test de comprensión lectora debería indicar siempre que el texto ha sido entendido por los alumnos, al igual que una prueba de producción escrita debería reflejar el nivel de competencia del discente en esta destreza<sup>5</sup>.

Así, una prueba final de evaluación elaborada exclusivamente a partir de preguntas referentes a micro-capacidades lingüísticas sería un test de diagnóstico o de referencia de los progresos realizados y, desde nuestro punto de vista, se correspondería en contenido y criterios de evaluación a los ejercicios de clase de realización periódica o a pruebas parciales de evaluación.

El principal problema que presentan las pruebas de micro-capacidades es la falta de correspondencia entre el uso que se hace de las mismas y los propósitos y tipos de evaluación para los que se utilizan, es decir, la tendencia de asumirlas como casi el único tipo de criterio evaluador de todo el proceso de enseñanza/aprendizaje. Si reconocemos

que la enseñanza, aprendizaje y evaluación de las micro-destrezas no suponen un fin en sí mismo, sino un medio para la adquisición de habilidades comunicativas globales, recomendamos un equilibrio entre preguntas pertenecientes a los diferentes niveles de las micro/macro-capacidades comunicativas en la realización de estas pruebas.

Asimismo, y según la opinión de Hutchinson y Waters (1987: 147), en la elaboración de estas pruebas de evaluación deberían seguirse siempre unos principios básicos:

Test what you can reasonably assume the learners have learnt (...); Your test should test what you actually want it to test (...); Don't, for example, write test items which demand specialist subject knowledge or cultural knowledge (...)

Compartimos la opinión de que el contenido de las pruebas de evaluación de las micro-capacidades debe coincidir con lo enseñado en clase y siempre de acuerdo a lo programado en la asignatura. Con evaluación de lo enseñado en clase no nos referimos, sin embargo, a la obligatoriedad de que todo lo impartido deba ser evaluado. Además de ser operativamente imposible, coincidiendo con Bachman y Palmer (1996), opinamos que el contenido de las pruebas de evaluación debe ajustarse a los objetivos del curso, es decir, a lo que se considere que el alumno debe haber aprendido: "Where the test takers are students in a language course, it is frequently necessary to develop a test that will provide feedback on how well students have learned the content of the course". En un contexto como el nuestro, inglés para fines específicos como idioma extranjero, y debido al nivel de competencia lingüística que en ocasiones presentan los discentes, nos vemos obligados a incluir en nuestra programación contenidos con el único propósito de servir de base al aprendizaje de habilidades comunicativas específicas. Con frecuencia sentimos la necesidad incluso de introducir explicaciones que no estaban recogidas en la programación original -estructuras gramaticales básicas, relaciones sintácticas, pronunciación, etc.- contenido que no tiene por qué aparecer en las pruebas de evaluación.

Un punto problemático en las consideraciones anteriores de Hutchinson y Waters (1987: 147) es el referente a la necesidad de evitar en el contenido de las pruebas material que pertenezca al área de especialización de los discentes<sup>6</sup>. La dificultad que entraña que el profesor conozca de antemano lo que pertenece o no al área de conocimiento del alumnado -tanto en cursos de inglés general como específico- haría de esta premisa una consideración idealmente teórica pero de dudosa aplicación práctica. Pensamos que sería imposible eliminar la proyección del conocimiento específico de los discentes en textos, por ejemplo, pertenecientes al entorno en el que se desenvuelven, es decir, tendencias musicales, deportes, moda, medios tecnológicos, etc.

Sin embargo, creemos que la frase de Hutchinson y Waters "Items which demand specialist subject knowledge" se circunscribe especialmente a ámbitos del inglés específico. Evitar contenido específico en pruebas de evaluación en IFE pondría de manifiesto una falta de correspondencia entre el contenido de la asignatura misma y la evaluación de los contenidos. Si el contenido de las pruebas de evaluación de las micro-capacidades debe ajustarse principalmente a lo enseñado en clase, debería existir de igual forma cierta coincidencia entre el tipo de material utilizado en ambas etapas del

aprendizaje -general, técnico, específico, entre otros-. Más que contenido específico en pruebas de evaluación, lo que debería evitarse es la elaboración de preguntas o ejercicios que puedan realizarse utilizando exclusivamente el conocimiento específico del discente: preguntas de comprensión lectora cuyas respuestas dependan exclusivamente del conocimiento específico, traducción de vocablos altamente especializados, elaboración de diagramas, tablas, etc. en pruebas de producción escrita donde no se evalúe el nivel de competencia en esta destreza, buscar sinónimos, etc.

Pensamos, no obstante, que en el caso de los tests de evaluación de las micro-capacidades, tests de diagnóstico, ejercicios periódicos de clase y similares (no considerados como un fin en sí mismos) podría utilizarse material diferente al incluido en el contenido específico del curso. Mientras que la utilización esporádica en clase de textos, videos y ejercicios pertenecientes a material de inglés general u otras especialidades suele acogerse de forma positiva por el alumnado, hemos observado una reacción de oposición al utilizar en pruebas de evaluación, sea cual fuere el propósito de las mismas, material distinto al programado en la asignatura.

No cabe duda de que las pruebas de evaluación de la competencia específica, por su adecuación al contenido de los diferentes apartados del programa, se circunscriben principalmente a la enseñanza de la lengua a nivel académico, “Language instruction domain situations in which language is used for the purpose of teaching and learning of language” (Bachman y Palmer 1996: 43), máxime cuando hay casos -principalmente cuando el uso de la lengua inglesa se limita al entorno académico- en los que resulta difícil identificar los usos lingüísticos (*TLU*) fuera del aula, es decir, la utilización de la lengua en ámbitos profesionales.

Destacaríamos una mayor necesidad de utilizar este tipo de pruebas de forma periódica con alumnos universitarios de IFE de primer curso que en cursos posteriores. La conveniencia de utilizar en alumnos de primer curso pruebas de evaluación correspondientes a la enseñanza/aprendizaje de la lengua inglesa a nivel académico -evaluación de habilidades lingüísticas de acuerdo con el contenido de la asignatura- vendría determinada por diferentes razones, entre las que destacamos:

- Una mayor dificultad por parte del alumnado de primer curso para identificar el uso real de la lengua fuera del aula (usos comunicativos específicos).
- Adecuación de las pruebas al contenido de las asignaturas. Teniendo en cuenta las características personales, opinión y preferencias, nivel de competencia lingüística, etc., incluimos en las asignaturas de primer curso un contenido mayor de aspectos pertenecientes al nivel de las micro-capacidades.
- Necesidad de una evaluación de diagnóstico continua debido a la heterogeneidad en el nivel de competencia lingüística de los discentes.

No obstante, en pruebas de evaluación de las micro-capacidades lingüísticas en cursos en los que los usos comunicativos hayan sido claramente identificados y definidos, recomendamos una adecuada correlación entre el contenido del curso a nivel académico, los usos del inglés fuera del aula y las pruebas de evaluación. Encontramos en Bachman y Palmer (1996: 291-297) un claro ejemplo de coincidencia entre el

contenido de un curso de IFE a nivel académico, el uso real de la lengua en ámbitos profesionales y el contenido de las pruebas de evaluación. El propósito de la prueba “Partially developed example of a syllabus-based diagnostic achievement test for students in an ESP program” recogida en el libro de Bachman y Palmer, consiste en corroborar si los discentes (alumnos no nativos con nivel intermedio en lengua inglesa) de un curso de inglés para trabajadores de una compañía telefónica han asimilado las habilidades programadas en los objetivos del curso. El test se hizo para diagnosticar el aprendizaje de estructuras y signos de puntuación en la escritura de una carta formal. La adecuación entre las pruebas de evaluación con el uso real de la lengua en contextos específicos es siempre una práctica recomendable:

It is our view that rather than automatically using language instructional tasks as the starting point for developing language test tasks, we must first determine the degree to which the instructional tasks correspond to the real-life tasks that the students will encounter after leaving the instructional setting. (Bachman y Palmer 1996: 105).

#### 4. PRUEBAS DE EVALUACIÓN DE CAPACIDADES COMUNICATIVAS (MACRO-CAPACIDADES)

Definimos capacidad comunicativa como la unión existente entre el conocimiento de la lengua y la habilidad de poner en práctica dichos conocimientos en usos específicos, según Savignon (1983):

Dynamic rather than static (...) It depends on the negotiation of meaning between two or more persons (...) It is context specific. Communication takes place in an infinite variety of situations, and success in a particular role depends on one's understanding of the context and on prior experience of a similar kind.

En una primera aproximación a las pruebas de evaluación de macro-destrezas comunicativas, coincidimos con la definición de Hutchinson y Waters (1987: 150) en cuanto al contenido y propósitos de las mismas:

The move towards proficiency testing fits very neatly with the concept of ESP which is crucially concerned with enabling learners to perform certain language tasks. Proficiency tests for specific purposes should, therefore, be able to give a reliable indication of whether a candidate is proficient enough to carry out the tasks that will be required.

Considerando las palabras de Hutchinson y Waters, el concepto de los tests de evaluación de la competencia lingüística (tests de macro-destrezas comunicativas) se ajustaría a los objetivos mismos del inglés específico en cuanto a la capacidad de evaluar la competencia de los discentes para realizar actos comunicativos concretos.

En la evaluación del nivel de la competencia lingüística podrían utilizarse escalas descriptivas de puntuación referentes al nivel de conocimiento adquirido por los discentes.



tes en la asimilación de los diferentes usos comunicativos, escalas como la utilizada por el British Council y recogida en Hutchinson y Waters (1987: 150): *Expert user; Very good user; Good user; Competent user; Modest user; Limited user; Extremely limited user; Intermittent user and Non-user*, o la utilizada por el ACTFL (American Council for the Teaching of Foreign Languages) con los siguientes descriptores: *Intermediate-high, Intermediate-mid, Intermediate low, Novice-High, Novice-Mid, Novice-low* en Hughes (1989: 89-90).

Adecuadas siempre a los propósitos de la prueba y al tipo de destreza a evaluar, habría que vigilar el número de subdivisiones incluidas en la escala, ya que descripciones demasiado o poco detalladas podrían presentar problemas a la hora de ubicar a los candidatos en la banda de puntuación correspondiente. Una división en seis niveles de conocimiento: *Inicial, Elemental, Elemental/Alto, Intermedio, Intermedio/Alto, Avanzado*, y donde la banda *Intermedio* representara el nivel mínimo de competencia lingüística aceptada, es decir, la capacidad para realizar actos comunicativos no complejos de forma aceptable, podría ser adecuada para nuestros propósitos.

The Interagency Language Roundtable (ILR) description of 'limited working proficiency' (the ILR scales relate to the use of language for a specific purpose, being used for assessing the foreign language ability of US Government agency staff) is as follows: Able to satisfy routine social demands and limited work requirements. Can handle with confidence but not with facility most social situations including instructions and casual conversations. (Hughes 1989: 104)

Sin embargo, pensamos que uno de los principales problemas que presenta la aplicación de los test de macro-capacidades reside en el diseño de pruebas que evalúen de forma efectiva el nivel de competencia comunicativa global de los discentes, en lugar de evaluar una serie de partes o contenidos lingüísticos relacionados entre sí (Bachman y Palmer 1982; Oller 1983).

Hutchinson y Waters (1987: 150) también se manifiestan en el mismo sentido:

Proficiency tests are primarily criterion-referenced. What should the criteria be? Should they vary with different subject areas? What skills and knowledge enable someone to perform particular tasks?

La dificultad de identificar las habilidades lingüísticas subyacentes que permiten llevar a cabo cada uno de los usos específicos, junto con la consideración de que el nivel de competencia de los discentes no debe basarse en el nivel medio del grupo sino en la realización satisfactoria de un acto comunicativo concreto, serían otros de los inconvenientes que presentaría la evaluación del nivel de competencia lingüística. En este contexto, Carroll (1993: 8) subraya dos aspectos importantes con respecto a la evaluación de los usos comunicativos específicos: "The individual must understand what sort of result is to be achieved and the individual needs to have some idea of the criteria by which performance will be assessed".

Asimismo, de acuerdo con Bachman y Palmer (1996: 44), si definimos *task* como la actividad o conjunto de actividades que se desarrollan en un contexto determinado con

la finalidad de alcanzar unos objetivos comunicativos específicos, otro de los problemas subyacentes en la aplicación de pruebas globales de evaluación sería la dificultad de identificar dichos usos comunicativos. La identificación de los usos específicos en IFE -tanto a nivel académico como profesional, *real-life tasks*- juega un papel decisivo en la elaboración y administración de las pruebas de macro-destrezas comunicativas, ya que consideramos necesario mantener una correspondencia directa entre los usos comunicativos específicos y el contenido de las diferentes pruebas de evaluación: "If the tasks in the language instructional domain correspond closely to those in an appropriate real-life domain, then the test developer can use tasks in either domain or both as a basis for developing test tasks" (Bachman y Palmer 1996: 104).

Así, cuando los usos comunicativos de un grupo de discentes estén claramente definidos (mantener una conversación telefónica, por ejemplo), recomendamos la utilización de pruebas que evalúen el nivel de competencia global en tales usos. De esta forma, además de un conocimiento lingüístico-gramatical, por ejemplo, se evaluaría la forma en la que la lengua se utiliza para conseguir una comunicación efectiva y la capacidad de utilizarla en su contexto. Asimismo, si los objetivos de la enseñanza/aprendizaje a nivel académico coinciden claramente con los usos de la lengua fuera del aula, las pruebas de evaluación de las micro-capacidades y las pruebas globales de evaluación ofrecerían muchas coincidencias:

Proficiency whether or not the specific abilities measured by a given proficiency test actually differ from those measured by a given achievement test will depend, of course, on the extent to which the theory upon the proficiency test is based differs from that upon which the syllabus is based (Bachman 1990: 71).

Por otra parte, y en los casos en los que los usos específicos del grupo no estén claramente definidos, las pruebas de evaluación deberían evaluar el nivel de competencia alcanzado en la asimilación de los contenidos del curso, tratando de maximizar siempre la relación entre el inglés académico programado (*syllabus-based tests*) y los posibles usos comunicativos (*real-life tasks*). Pensamos que la dificultad para determinar los usos de la lengua fuera del aula no nos debe llevar en cursos de IFE a crear pruebas de evaluación basadas exclusivamente en puntos aislados del programa.

Concluimos reiterando la necesidad de que exista una relación de reciprocidad entre los procesos de enseñanza/aprendizaje de los contenidos que conforman la programación de lengua inglesa en cursos de IFE y la evaluación de los mismos. Además, afirmamos que si la finalidad última de las pruebas de evaluación se basa en recabar información sobre el nivel de competencia lingüística de los discentes con el objetivo de tomar decisiones a partir de las mismas (diagnosticar carencias específicas, pasar al punto siguiente del programa, aprobar un examen, conseguir un puesto de trabajo, etc.), la administración de un tipo u otro de prueba de evaluación vendrá determinada por el tipo de decisión a tomar. Destacamos, en nuestro contexto, la importancia de las pruebas de evaluación de las micro-capacidades lingüísticas como pruebas parciales y de diagnóstico, y de las pruebas de evaluación de las macro-capacidades comunicativas como medio para evaluar el nivel de competencia lingüística global. La evaluación de actos

comunicativos específicos convertirá a este último tipo de test en una de las pruebas más importantes en IFE, tanto a nivel académico como en ámbitos profesionales.

## NOTAS

1. Para una mayor comprensión de los principios básicos de la evaluación de idiomas y la elaboración de pruebas de evaluación, véase Alderson, Clapham, y Wall (1995) y Weir (1993). Asimismo, el diccionario publicado por Davies, Brown, Elder, Hill, Lumley y McNamara (1999) presenta de forma concisa los conceptos más importantes que subyacen en los procesos de evaluación.
2. Basado en la experiencia práctica de sus autores, en Weir y Roberts (1994) encontramos una revisión teórica acompañada de una exhaustiva exposición de tipos de pruebas, proporcionando una guía esencial de los procedimientos que concurren en la elaboración y administración de pruebas de evaluación en contextos diferentes.
3. Para tener una visión general de los diferentes acercamientos metodológicos aplicados a la interpretación de resultados en pruebas de evaluación en contextos generales y específicos (*criterion-reference, generalizability theory, IRT, etc.*), véase Bachman (1990: 160-233). Los artículos de Brown (1995), Elder y Brown (1997) y McNamara (1990) profundizan en la proyección de los acercamientos metodológicos anteriores.
4. Para profundizar en el concepto de *Validación* -necesidad de que los tests se adecuen a los propósitos y fines de la evaluación-, consúltese el libro editado por Alderson y North (1991).
5. Weir (1998) analiza, mediante ejemplos prácticos, la necesidad de elaborar y seleccionar pruebas de comprensión lectora en ámbitos académicos que evalúen el nivel global de competencia lingüística de los discentes.
6. En Clapham (1996) y McNamara (1997) se nos presentan estudios sobre los diferentes factores (tipo de tests, características de los discentes, conocimiento específico, etc.) que pueden incidir en el desarrollo y los resultados obtenidos en pruebas de evaluación en IFE. La incidencia ejercida por estos factores representa una de las líneas de investigación más estudiada en los últimos años. Destacamos asimismo el libro editado por Bachman y Cohen (1998) donde se profundiza en la aplicación de nuevos enfoques en la evaluación de idiomas: "testing cross-cultural pragmatics". Los procesos de evaluación se analizan de forma conjunta con otras áreas de la lingüística aplicada como la sociolingüística y la pragmática.

## BIBLIOGRAFÍA

- Alderson, J.C. y B. North, eds. 1991. *Language Testing in the 1990s: the communicative Legacy*. Londres: Macmillan.
- Alderson, J.C., C. Clapham, y D. Wall 1995. *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- Alderson, J.C. y A.H. Hughes, eds. 1981. "Issues in Language Testing". *ELT Documents III*. Oxford Pergamon Press.
- Bachman, L.F. 1990. *Fundamental Consideration in Language Testing*. Oxford: Oxford University Press.
- Bachman, L.F. y A.D. Cohen, eds. 1998. *Interfaces between Second Language Acquisition and Language Testing Research*. Cambridge: Cambridge University Press.

- Bachman, L.F. y A. Palmer. 1982. "The construct validation of some components of communicative proficiency". *TESOL Quarterly* 16, 4: 449-65.
- Bachman, L.F. y A. Palmer. 1996. *Language Testing Practice*. Oxford: Oxford University Press.
- Brown, A. 1995. "The Effect of Rater Variables in the Development of an Occupational Specific Language Performance Test". *Language Testing* 12, 1: 1-15.
- Carroll, J.B. 1993. *Human Cognitive Abilities: A Survey of Factor Analytic Studies*. Cambridge: Cambridge University Press.
- Clapham, C. 1996. *The Development of IELTS: A Study of the Effect of Background Knowledge on Reading Comprehension*. Cambridge: Cambridge University Press.
- Candlin, C.B., J.H. Leather y C.J. Bruton. 1976. "Doctors in Casualty: Applying Communicative Competence to Components of Specialist Course Design". *International Review of Applied Linguistic* 14: 246-72.
- Davies, A., A. Brown, C. Elder, K. Hill, T. Lumley y T. McNamara. 1999. *Dictionary of Language Testing*. Cambridge: Cambridge University Press.
- Davies, S. y R. West. 1984. *The Pitman Guide to English Language Examinations*. Londres: Pitman Publishing Ltd.
- Elder, C. y A. Brown. 1997. "Performance Testing for the Professions: Language Proficiency or Strategic Competence?". *Melbourne Papers in Language Testing* 6, 1.
- Hughes, A. 1989. *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- Hutchison, T. y A. Waters. 1987. *English for Specific Purposes*. Cambridge: Cambridge University Press.
- Mumby, J. 1978. *Communicative Syllabus Design*. Cambridge: Cambridge University Press.
- McNamara, T. 1990. "Item Response Theory and the Validation of an ESP Test for Health Professionals". *Language Testing* 7, 1: 52-76.
- McNamara, T. 1997. "Problematising content Validity: the Occupational English test as a measure of Medical Communication". *Melbourne Papers in Language Testing* 6, 1.
- Oller, J. W., ed. 1983. *Issues in Language Testing Research*. Rowley, MA.: Newbury House.
- Savignon, S. 1983. *Communicative Competence: Theory and Classroom Practice*. Nueva York: Addison-Wesley.
- Weir, C. 1993. *Understanding and Developing Language Tests*. Londres: Prentice Hall.
- Weir, C. y Roberts, J. 1994. *Evaluation in ELT*. Oxford: Blackwell.
- Weir, C. 1998. "The Selection of Texts and Tasks for Testing and Teaching Academic Reading Ability in English". *Quality in Learning in English Medium Higher Education*. Ankara: Bilkent University Press.