

03070

1 lej.

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

UNIDAD ACADÉMICA DE LOS CICLOS  
PROFESIONAL Y DE POSGRADO DEL  
COLEGIO DE CIENCIAS Y HUMANIDADES  
CENTRO DE ENSEÑANZA DE LENGUAS EXTRANJERAS

UN SISTEMA DE PREPARACION PARA CANDIDATOS A BECAS  
PARA ESTUDIAR EN PAISES DE HABLA INGLESA:  
LA EVALUACION BASADA EN NECESIDADES

T E S I S

COMO REQUISITO PARA LA OBTENCION DEL GRADO DE  
MAESTRIA EN LINGUISTICA APLICADA

ANN HILDRETH

MEXICO, D.F.

1983

TESIS CON  
FALLA DE ORIGEN



Universidad Nacional  
Autónoma de México



## **UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso**

### **DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

## INDICE

Capítulo I - Introducción a la evaluación basada en necesidades.	1
1.0 Introducción general	1
1.1 Esbozo del presente trabajo	3
1.2 El análisis de necesidades: su importancia actual	4
1.3 La evaluación basada en necesidades	6
1.4 El lugar del análisis de necesidades en el proceso de diseño de cursos y de instrumentos de evaluación.	7
Capítulo II - Definición de un modelo para el análisis de necesidades	11
2.0 Introducción	11
2.1 Descripción de varios modelos	11
2.1.1 El modelo de Strevens	11
2.1.2 El proceso descrito por MacKay	13
2.1.3 Brumfit	14
2.1.4 Shaw	15
2.1.5 El Consejo de Europa	16
2.1.6 El modelo de Munby	18
2.2 Discusión de las ventajas y desventajas de los modelos presentados	19
2.2.1 Los modelos de Strevens y Munby	
2.2.1.1 La taxonomía de cursos para propósitos específicos	20
2.2.1.2 El estudio del contenido lingüístico	20
2.2.1.3 La definición de objetivos en términos comunicativos	21
2.2.1.4 Problemas internos en el modelo de Munby	22
2.2.1.5 Otros niveles	23
2.2.2 El Consejo de Europa	24
2.2.3 Contribuciones de MacKay, Brumfit y Shaw	24
2.3 El modelo de análisis del presente trabajo	25

Capítulo III	Los resultados del estudio de necesidades	28
3.0	Introducción	28
3.1	Justificación - necesidades nacionales e institucionales	29
3.2	Evaluación	
3.2.1	La evaluación del nivel de la población meta	29
3.2.2	Información acerca de cursos ya existentes	31
3.3	Perfil del participante	34
Capítulo IV	Definición y discusión de conceptos básicos en la evaluación	40
4.0	Introducción	40
4.1	Tipos de reactivos y exámenes	43
4.1.1	Reactivos de puntos aislados versus exámenes integrativos	43
4.1.2	Reactivos o exámenes objetivos y subjetivos	48
4.1.3	Exámenes de aprovechamiento	52
4.1.4	Examen de requisito	53
4.1.5	Examen de dominio	53
4.1.6	Examen comunicativo	55
4.1.7	Examen de actuación versus examen de dominio comunicativo	58
4.2	Evaluación por criterios versus evaluación por la norma	60
4.3	Confiabilidad	66
4.4	Validez	66
4.5	¿Qué evaluar? Consideraciones sobre la validez de construcción	73
4.5.1	Estudios de la construcción "Dominio lingüístico"	73
4.5.2	Estudios de la construcción "Competencia comunicativa"	76
Capítulo V	El examen de requisito para un sistema de preparación para becarios	89
5.0	Introducción	89
5.1	Evaluación del nivel lingüístico	92
5.2	Evaluación del nivel comunicativo	94
5.2.1	Micro-habilidades	95
5.2.2	Tareas comunicativas globales	96
5.3	Practicabilidad	98
5.4	El diseño del nivel comunicativo	99
5.4.1	Eventos y actividades comunicativas	99
5.4.2	Los parámetros de evaluación y niveles de actuación	99
5.4.2.1	Los parámetros de evaluación	99
5.4.2.2	La especificación de los tres niveles de actuación para las cuatro fases del examen	102

5.5	La especificación de escalas de evaluación para el nivel comunicativo	103
5.6	La especificación de las micro-habilidades a medir, con apuntes acerca de cómo se podrían medir	116
5.7	Diseño de tareas de comprensión y producción	130
5.7.1	Tareas de comprensión	131
5.7.2	Tareas de producción	132
Capítulo VI Ejemplos y conclusiones		135
6.0	Introducción	135
6.1	Los ejemplos	135
6.1.1	Introducción al texto	136
6.1.2	La prueba de micro-habilidades	139
6.1.2.1	Los reactivos	139
6.1.2.2	Análisis de los reactivos en términos de las micro-habilidades del nivel 2.	142
6.1.3	La prueba de actuación	143
6.1.3.1	Instrucciones para la prueba de actuación	143
6.1.3.2	Comentarios sobre la prueba de actuación	143
6.2	Comentarios sobre las dos pruebas	143
6.3	Conclusiones sobre la evaluación basada en necesidades	144

## INDICE DE CUADROS

1-1	Modelo de diseño de cursos de Candlin, Kirkwood y Moore	8
1-2	Modelo de la evaluación basada en necesidades: modificación del modelo de Candlin, Kirkwood y Moore	9
2-1	Resumen de información a recabar en el análisis de necesidades - Consejo de Europa	17
2-2	Modificación del procesador de necesidades comunicativas del modelo de Munby	23
4-1	Matriz para un estudio convergente-discriminante	78
5-1	Diseño global del examen de requisito para becarios	91
5-2	Parámetros de evaluación y nivel meta del modelo de Munby	100
5-3	Especificación de niveles de actuación: Comprensión auditiva (conferencias)	104
5-4	Especificación de niveles de actuación: Comprensión de lectura	107
5-5	Especificación de niveles de actuación: Producción escrita	110
5-6	Especificación de niveles de actuación: Interacción oral (seminario)	113
5-7	Micro-habilidades a medir: Comprensión auditiva (conferencias)	117
5-8	Micro-habilidades a medir: Comprensión de lectura	120
5-9	Micro-habilidades a medir: Producción escrita	125
5-10	Micro-habilidades a medir: Interacción oral (seminario)	127
5-11	Referencia de las dimensiones en la comprensión de lectura	131
5-12	Parámetros generales y específicos	133

CAPITULO I  
INTRODUCCION A LA EVALUACION  
BASADA EN NECESIDADES

1.0 Introducción general

El propósito del presente trabajo es proponer un esquema de la forma que podría tomar un examen para candidatos a becas para estudiar en países de habla inglesa. Es importante que las instituciones que otorgan becas traten de asegurar que las personas que logran obtenerlas tengan las mayores posibilidades de tener éxito en sus estudios en el extranjero. Dominar la lengua es sólo un factor entre varios que influyen en el éxito o no, pero consideramos que es un factor importante, y es el factor que el lingüista aplicado puede tratar de medir.\*

Aquí intentamos proponer un instrumento de evaluación que sea válido y confiable, en el mayor grado posible, para medir la capacidad del candidato de enfrentar la situación de hacer estudios de posgrado en países de habla inglesa.

---

\* Factores como personalidad, el hecho de tener o no a familiares en la nueva situación, etc. influyen también. Se ha estudiado también la influencia de las modalidades de estudio de los becarios en su desarrollo académico.<sup>1</sup>

Mientras que en este trabajo se trata esencialmente la evaluación, se considera que esta evaluación debería formar parte de un sistema global de preparación y evaluación, en que los candidatos a becas no sólo tendrían que presentar un examen, sino que tendrían también la oportunidad de prepararse en cur sos especialmente diseñados para ellos.\*

El desarrollo del esquema de este examen se basa en dos áreas; el análisis de las necesidades comunicativas de los examinados, y la evaluación en sí. Se pretende presentar el diseño de un examen que toma en cuenta de manera sistemática, las ne cesidades comunicativas de los candidatos a becas, y que refleja una serie de decisiones tomadas en base a la consideración de algunos problemas claves en la evaluación.

Mientras que se presentan en el capítulo VI ejemplos de dos subpruebas, no se pretende desarrollar aquí un examen completo. La contribución del trabajo yace en la especificación de los aspectos a evaluar en un examen de becario en particular, y en general en la presentación de una manera de diseñar exámenes basados en criterios, y tomando en cuenta sistemáticamente las necesidades comunicativas del examinando.

---

\*Este trabajo tiene una relación estrecha con uno del Mtro. David Howard, que se ocupa del diseño de un curso de este tipo.

## 1.1 Esbozo del presente trabajo

En este primer capítulo se introduce el tema, delimitando el trabajo, e introduciendo el concepto de la evaluación basada en necesidades.

En el capítulo II, se examinan algunos modelos para el análisis de necesidades, viendo la manera en que pueden contribuir a un modelo adecuado para el estudio de necesidades de becarios mexicanos. En el capítulo III, se presentan, en forma resumida, los resultados de un estudio de necesidades basado en este modelo.\* Estos resultados luego sirven para el diseño del examen.

En el capítulo IV, se examinan algunos conceptos de la evaluación que repercuten en el diseño del examen de requisito para becarios. Estos conceptos se discuten con relación a las decisiones acerca de la forma que este examen debe tomar.

El diseño del examen de requisito para becarios aparece en el capítulo V, con especificaciones detalladas para el nivel comunicativo. Las conclusiones, y algunos ejemplos del examen comunicativo (en el área de comprensión de lectura) aparecen en el último capítulo.

---

\* Esta información fue recabada en su mayor parte en 1981, dentro del proyecto llevado a cabo por Colle, Hildreth y Howard, para la materia Diseño de Cursos, de la Maestría en Lingüística Aplicada.

## 1.2 El análisis de necesidades: su importancia actual.

La importancia del análisis de necesidades en el diseño de cursos es bien conocida. La idea que también es paso indispensable en el diseño de los instrumentos de evaluación está ganando aceptación. Este énfasis en el análisis de necesidades se relaciona con el surgimiento de un acercamiento al diseño de cursos que toma unidades comunicativas (funciones, nociones, habilidades, etc.) como base de su organización.

Allen y Widdowson notan una tendencia general en que el inglés ha ganado importancia en la educación superior, en países que no son de habla inglesa. Ellos señalan que debido a esta tendencia, la enseñanza del inglés ha cambiado, dejando su orientación hacia las habilidades de leer, escribir, hablar y entender, y adoptando como metas "...la capacidad básica de usar la lengua para comprender y (en menor grado) comunicar información asociada con sus estudios especializados."<sup>1</sup> \*

Mientras que antes se veía al inglés como una disciplina, una materia aislada e independiente, el cambio en el papel del inglés en la educación superior ha producido una nueva perspectiva en que se percibe el inglés como instrumento de comunicación de contenidos relacionados con otras áreas. Este

---

\* Notaríamos que Allen y Widdowson aquí presentan una visión muy simplificada del uso de la lengua, omitiendo toda una gama de funciones importantes en el uso de la lengua en las disciplinas académicas.

cambio, entonces, ha producido un énfasis en el uso comunicativo de la lengua. En el momento en que se concibe a la lengua principalmente como instrumento para uso con disciplinas diversas, surge la posibilidad de que diferentes disciplinas puedan presentar diferentes exigencias en cuanto al uso de la lengua. Se encuentra, entonces, una diversificación de metas en cursos de lengua para propósitos específicos, y el surgimiento del análisis de necesidades para determinar cuáles son las exigencias y características particulares del uso de la lengua en grupos particulares.

Según el punto de vista de Allen y Widdowson, la importancia que se ha dado al análisis de necesidades viene en gran parte de un cambio en el lugar del inglés (o, podemos añadir, de un creciente reconocimiento de su importancia) en los estudios superiores en diferentes partes del mundo.

Desde otra perspectiva, Keith Johnson parece ver el análisis de necesidades como consecuencia de la adopción de un enfoque funcional-nocional por otras razones. El señala que un inventario completo de todas las nociones y funciones que un individuo pudiera un día utilizar, sería demasiado largo, y totalmente inadecuado para propósitos de enseñanza.<sup>2</sup> Para llegar a una lista que pudiera aplicarse a la enseñanza, se necesitaría de un criterio para incluir unos elementos, y excluir otros. El criterio sugerido es la utilidad para el grupo específico.

Cuando se utilizan estructuras como las unidades a enseñar, los criterios básicos para la inclusión o exclusión de uno u otro elementos han sido una lógica interna de graduación de estructuras, tomando en cuenta su complejidad, su supuesta dificultad para aprender, y tal vez su frecuencia. La aplicación de estos criterios no requiere de un análisis de necesidades de grupos particulares: la aplicación de un criterio de utilidad de unidades comunicativas sí lo requiere. Johnson, entonces, ve el análisis de necesidades como consecuencia práctica de la adopción de nociones, funciones y otras unidades comunicativas como unidades para la enseñanza.

### 1.3 La evaluación basada en necesidades

Así como es reconocida la importancia del análisis de necesidades para el diseño de cursos, también su importancia en el diseño de instrumentos de evaluación se comienza a reconocer. Este reconocimiento es consecuencia de dos fenómenos. Por una parte, la existencia de cursos cuyos objetivos se han fijado con base en las necesidades comunicativas de los participantes, ha ocasionado el desarrollo de exámenes de aprovechamiento basados en estos mismos objetivos. Es decir, la evaluación relacionada con cursos basados en necesidades es evaluación también basada en necesidades.

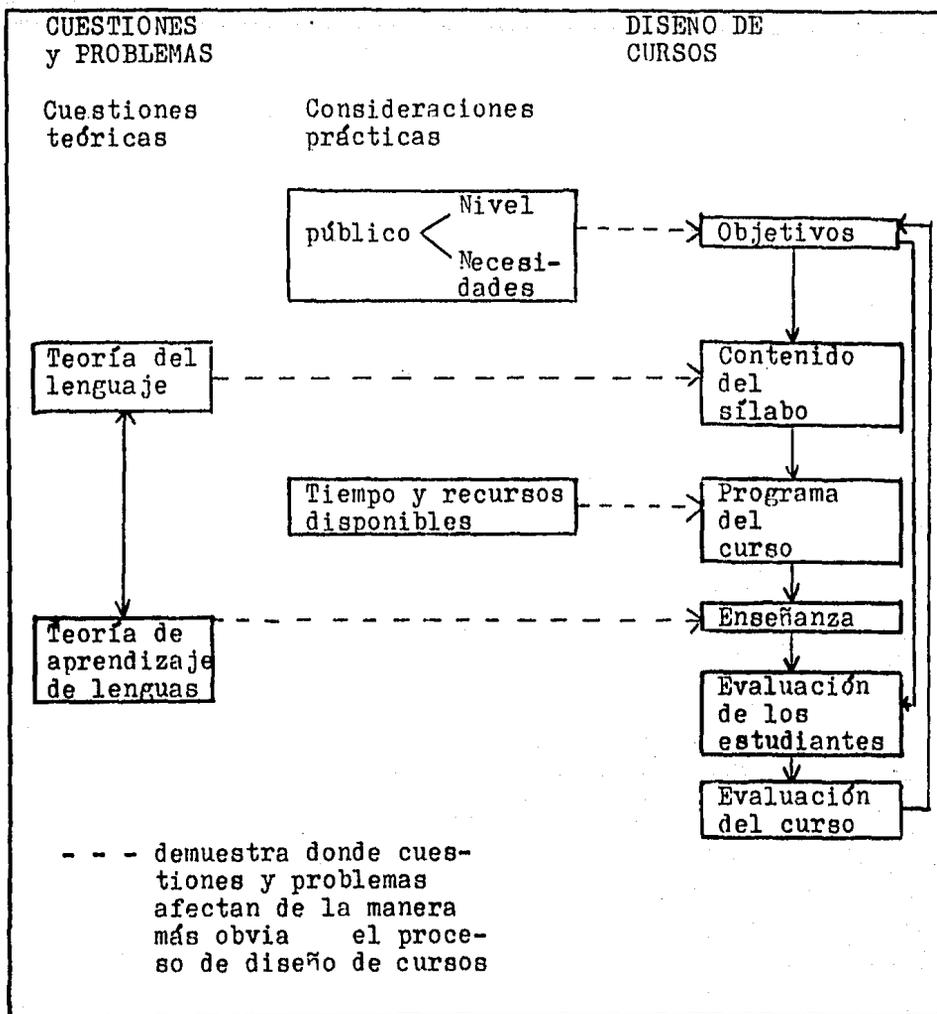
Por otra parte, otro factor ha alentado el desarrollo de exámenes basados en análisis de necesidades comunicativas, aun fuera del contexto de un curso específico. Este es el caso de

exámenes de requisito, exámenes que requieren instituciones, por ejemplo para recibir un diploma, ingresar a un trabajo, u obtener una beca. En estas situaciones, el interés de fundamentar los exámenes en un análisis de necesidades viene de un cambio en la percepción de lo que significa conocer una lengua, y por lo tanto, cierta desconfianza en la validez del concepto de "dominio general de una lengua". Esta desconfianza ha traído como consecuencia, la desconfianza en los exámenes de dominio general. Una discusión más a fondo de estas cuestiones se encuentra en la sección 4.5 " Qué evaluar? Consideraciones sobre la validez de construcción."

#### 1.4 El lugar del análisis de necesidades en el proceso de diseño de cursos y de instrumentos de evaluación

En su artículo "Study Skills in English: Theoretical Issues and Practical Problems" Candlin, Kirkwood y Moore presentan un modelo de diseño de cursos que parece resumir las relaciones entre los elementos esenciales que entran en este trabajo. En el cuadro 1-1, se reproduce el modelo tal como lo presentan los autores mencionados.

Este modelo nos parece una buena representación esquemática de los factores que han intervenido en este trabajo. No se le ha utilizado como "instructivo" a seguir en el diseño de cursos ni de exámenes, sino como un marco que refleja la interacción de diferentes elementos. El modelo tiene varias características que lo hacen especialmente adecuado como marco de referencia para el trabajo.



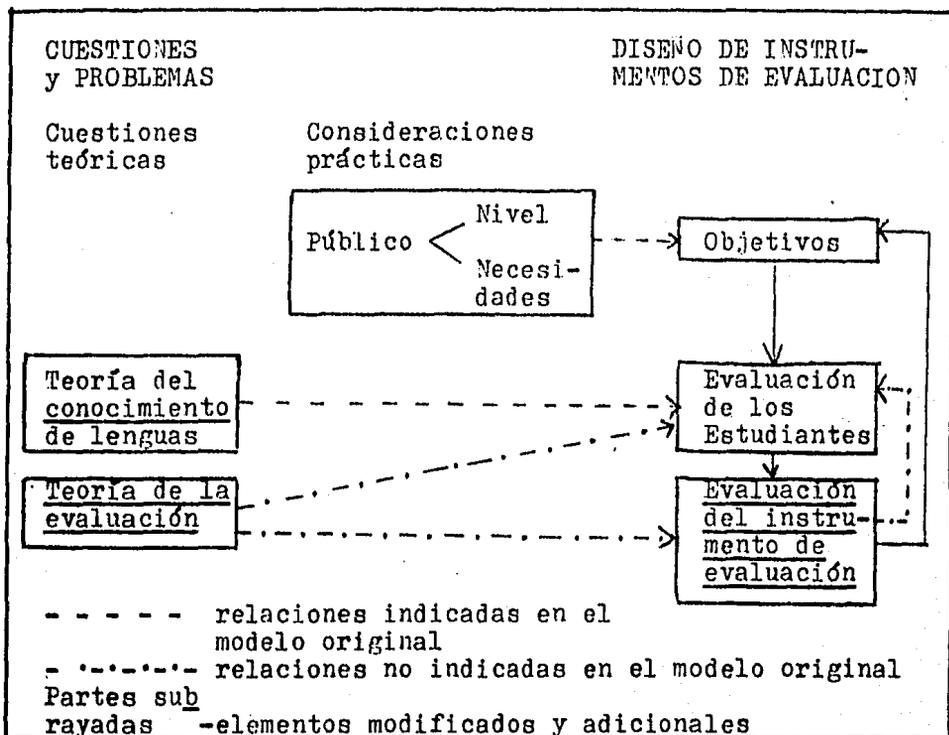
CUADRO 1-1

MODELO DE DISEÑO DE CURSOS DE CANDLIN, KIRKWOOD Y MOORE<sup>3</sup>

- i) los objetivos se derivan directamente de las necesidades y el nivel del público en un principio, y hay retroalimentación posterior basada en la evaluación del curso.
- ii) el contenido del sílabo (unidades a enseñar) - su naturaleza (funciones, nociones, estructuras, habilidades) y el contenido (cuáles funciones, nociones, estructuras y habilidades) se derivan de los objetivos, pero con la fuerte influencia de la teoría del lenguaje.

iii) la evaluación se deriva por una parte de la enseñanza, pero, por la otra, se deriva directamente de los objetivos. Esto evita que la evaluación se base demasiado en el contenido y procedimientos específicos del curso, permitiendo así una mejor evaluación de la manera en que el curso cumple los objetivos.

Otra ventaja de este modelo para el presente trabajo, es que permite, con algunas modificaciones, aislar el elemento de la evaluación del modelo global temporalmente, para desarrollar un examen de requisito que tendría los mismos objetivos de un curso, pero sin necesariamente desarrollar el curso con anterioridad. El cuadro 1-2 presenta el elemento de evaluación del modelo de Candlin, Kirkwood y Moore, aislado, y con algunas modificaciones.



CUADRO 1-2  
 MODELO DE LA EVALUACION BASADA EN NECESIDADES  
 MODIFICACION DEL MODELO DE CANDLIN, KIRKWOOD Y MOORE

Para llegar a este modelo aislado de evaluación se modificó el modelo original de Candlin et al., de la siguiente manera. Se omitieron los elementos que se refieren directamente al diseño de curso (contenido del sílabo, programa del curso, recursos disponibles, y enseñanza.) Se cambió "teoría del aprendizaje de lenguas" por "teoría del conocimiento de lenguas", y se agregó "teoría de la evaluación", elemento que influye tanto en la evaluación de los estudiantes, como en la evaluación del mismo instrumento de evaluación. Finalmente se agrega una indicación de la retroalimentación entre la evaluación del instrumento al mismo instrumento. Es decir, una indicación de la poca confiabilidad del instrumento, o de la baja validez, podría influir en la forma que toma el examen sin necesariamente implicar un cambio en los objetivos a medir.

Así modificado entonces, el cuadro 1-2 representa esquemáticamente las interrelaciones entre los elementos más importantes en el diseño de instrumentos de evaluación basados en necesidades.

---

Notas de referencia bibliográfica - Capítulo I

<sup>1</sup>J. P. B. Allen y H. G. Widdowson, "Teaching the Communicative Use of English," en English for Specific Purposes, eds., Ronald MacKay y Alan Mountford (London:Longman, 1978), p. 56.

<sup>2</sup>Keith Johnson, Communicative Syllabus Design and Methodology (Oxford: Pergamon Press)1982), p. 40.

<sup>3</sup>C.N. Candlin, et al., "Study Skills in English: Theoretical Issues and Practical Problems," en English for Specific Purposes, eds. MacKay y Mountford, p. 191.

CAPITULO II  
DEFINICION DE UN MODELO  
PARA EL ANALISIS DE NECESIDADES

2.0 Introducción

En este capítulo se revisan varios modelos para el análisis de necesidades, con el objeto de ver cómo pueden contribuir a un modelo adecuado para el análisis de necesidades de becarios mexicanos. Luego, se describe el modelo utilizado en este trabajo.

2.1. Descripción de varios modelos.

2.1.1 El modelo de Strevens

Strevens, en un capítulo titulado "La enseñanza del inglés para propósitos específicos" plantea un proceso y un producto para el análisis de necesidades.<sup>1</sup> El no utiliza el término "análisis de necesidades", pero establece tres aspectos vitales del diseño de cualquier curso particular, en lenguas para propósitos específicos. Primero establece una taxonomía de tipos de cursos que caben dentro de la categoría de cursos de lenguas para propósitos específicos.\* También incluye aquí las consideraciones pedagógicas del nivel de los alumnos y la cantidad (duración y frecuencia) de instrucción.

---

\* Esta taxonomía está incluida también en el modelo presentado en la obra Communicative Syllabus Design de John Munby.

El segundo aspecto que considera Strevens es el contenido lingüístico. Aunque no profundiza en este aspecto, es claro que implica un estudio cuidadoso lingüístico (gramática, léxico, pronunciación, estilo) para descubrir cuáles son los elementos lingüísticos que distinguen el registro a estudiar. El tercer aspecto a considerar es "el propósito comunicativo que el estudiante busca controlar."<sup>2</sup> Da un ejemplo en que indica los propósitos comunicativos del inglés para propósitos ocupacionales, para un técnico o un ingeniero en una fábrica. Indica las 'categorías comunicativas': I. generales (tipos de comunicación comunes a un rango amplio de lugares de trabajo técnico), II. sociales (comunicación social con jefes, compañeros, etc.) y III. específicas (tipos de comunicación y temas que son específicos a una industria o una fábrica particular). Estas categorías incluyen funciones como "recibir y comprender instrucciones" y "expresar acuerdo y desacuerdo", nociones como "medición" y "cantidad", y habilidades como "comprender diagramas".\* Strevens da ejemplos de cada área indicada, en los medios oral y escrito, a nivel productivo y receptivo.

Strevens llega a su lista de funciones y ejemplos por medio de la introspección; reconoce las limitaciones de este proceso cuando dice que los resultados reflejan "los propósitos

---

\* Strevens utiliza el término global de 'categorías comunicativas' para todos; reserva el término 'función' para generalidades como 'función interpersonal del lenguaje'.

comunicativos que el autor considera, tal vez no acertadamente, ser las necesidades de estos estudiantes."<sup>4</sup>

El trabajo de Strevens concretiza tres áreas de información previa necesaria para el diseño de cursos: tipo de curso, contenido lingüístico y categorías comunicativas. Así da una respuesta parcial a la pregunta, ¿qué forma debe tomar el análisis de necesidades?

### 2.1.2 El proceso descrito por MacKay

MacKay, en su artículo "Identifying the Nature of the Learner's Needs" describe el proceso de análisis de necesidades seguido antes de establecer cursos de inglés (comprensión de lectura) en la Facultad de Medicina Veterinaria y Zootecnia de la U.N. A.M.<sup>5</sup> Hace resaltar la importancia de dos factores en la investigación de necesidades.

Primero hace hincapié en la importancia de hacer un estudio formal, basado en entrevistas estructuradas y cuestionarios, para llegar a conclusiones más confiables. Ahí aboga por una metodología rigurosa. Segundo, insiste en la importancia de hacer una investigación imparcial que no presuponga la necesidad de un curso, sino que trate de establecer si existe o no esta necesidad. No trata de establecer objetivos antes de asegurar que es necesario el curso. En el caso específico que describe, la investigación tiene como meta encontrar si leer el inglés era realmente indispensable para que los

alumnos terminaran con éxito sus estudios en la facultad, o para que ejercieran como profesionistas. Si este fuera el caso, se habría establecido una necesidad no sólo individual, de cada estudiante, sino también institucional, una necesidad de la misma facultad. Sólo habiendo comprobado esta necesidad de leer textos en inglés, necesidad individual e institucional, se justificaría plenamente establecer un curso, o institucionalizar un examen de requisito.

Así MacKay establece un nivel de análisis de necesidades institucional que debe complementar el análisis a nivel individual.

### 2.1.3 Brumfit

En su artículo "'Communicative' Language Teaching: An Educational Perspective" Brumfit habla de la responsabilidad del maestro. El dice:

... los maestros, representando al sistema educativo, necesitan poder demostrar que lo que se está enseñando es deseable, directa o indirectamente, para el bien de la sociedad en general.<sup>6</sup>

Brumfit no se refiere aquí directamente a un análisis de necesidades, sino al rol del maestro en el proceso educativo. Sin embargo, esta afirmación parece ser también de importancia para quienes se encargan de diseñar cursos. Si el maestro, que generalmente recibe instrucciones acerca de qué enseñar es responsable por lo que enseña sería el caso aun más

con las personas que determinan lo que se enseña, y lo que se evalúa. Tomando esta línea de pensamiento, el análisis de necesidades debe contemplar un análisis a nivel de la sociedad, para asegurar que los objetivos establecidos realmente sean "deseables para el bien de la sociedad", como lo dice Brumfit.

En un insituto privado, se podría argüir que el maestro no necesita demostrar que lo que enseña es positivo para el bien de la sociedad. Sin embargo, una situación en que lo que se enseña fuera en detrimento de la sociedad o de un sector de la sociedad, podría surgir un problema de ética para el maestro involucrado.

#### 2.1.4 Shaw

Shaw indica un análisis de los factores internos y externos que influyen en el proceso de desarrollo de un curso.<sup>7</sup> Menciona no sólo el análisis de las necesidades y los deseos de los estudiantes, y el tipo de curso, sino también la política y las necesidades a nivel institucional, las características particulares del estudiante, del sistema educativo, del maestro, de la situación lingüística y económica, y de las políticas gubernamentales y actitud de la gente. Observa que el diseñador de cursos debe estudiar la situación particular con el mayor cuidado posible, para adecuar el curso a ella. Shaw presenta, entonces un acercamiento que incluye la consideración de una gama amplia de factores institucionales, lingüísticos, educativos sociales y políticos.

### 2.1.5 El Consejo de Europa

El proceso de identificación de necesidades desarrollado para usar en el sistema de aprendizaje-enseñanza de lenguas que contempla el Consejo de Europa, es único en la amplitud de la perspectiva que proporciona.<sup>8</sup> Aquí, la identificación de necesidades no se concibe como un paso previo al diseño de cursos, sino como un proceso continuo, integrado a los aspectos de objetivos, evaluación, planes de estudios y recursos. No sólo es un proceso que involucra a las instituciones que ofrecen cursos, sino incluye también a los individuos que estudian idiomas, las instituciones que recurren a institutos que ofrecen cursos (para entrenar a su personal, por ejemplo), y a la sociedad. Es la manera en que se llega a un compromiso entre las ideas de los estudiantes acerca de objetivos, evaluación, programas y recursos, y las ideas de las instituciones y la sociedad. Dicen Richterich y Chancerel:

La identificación [de necesidades] en un acercamiento que contempla el sistema, centrado en el estudiante, necesariamente consistiría en la construcción, con él de un proyecto de aprendizaje, y el descubrimiento con él, de la manera en que podría cumplirlo institucionalmente, y dentro de la sociedad.<sup>9</sup>

Este análisis de necesidades se centra en el estudiante, pero reconoce que todo el proceso no parte de ahí, sino que se relaciona con, y gira alrededor de él. La función del análisis de necesidades es de ayudar al estudiante a definir sus necesidades y a encontrar la manera de satisfacerlas dentro de un

Recabada por:	Acerca de:	Cuando:	Propósito:
(1) El estudiante	Si mismo	Antes del curso	Hacer explícitos: sus objetivos; los métodos y tipos de evaluación que prefiere y requiere; los recursos de que dispone.
(2) El estudiante	Si mismo	Durante el curso	Evaluar el curso en términos de (1)
(3) El estudiante	Instituciones de enseñanza	Antes del curso	Decidir cuál institución le conviene.
(4) Instituciones de enseñanza	Ellos mismos	Continuamente	Determinar cómo adaptarse a necesidades y deseos de los alumnos en cuanto a (1)
(5) Instituciones de enseñanza	El estudiante	Antes del curso	Determinar sus objetivos, para poder sugerirle su programa, y formas de evaluación.
(6) Instituciones de enseñanza	El estudiante	Durante el curso	Buscar maneras de adaptarse a las necesidades del estudiante, que están evolucionando.
(7) Instituciones que recurren a instituciones de enseñanza	Uso de lenguas extranjeras	Continuamente	Decidir para qué se necesita la lengua extranjera, para proponer programas y decidir cuántos recursos destinar al entrenamiento en lenguas extranjeras.
(8) Sociedades	Ellos mismos	Continuamente	Evaluar los conocimientos que tienen de, y el uso que le dan a las lenguas extranjeras sus ciudadanos; evaluar programas ya existentes.

CUADRO 2-1

RESUMEN DE INFORMACION A RECABAR  
ANALISIS DE NECESIDADES - CONSEJO DE EUROPA<sup>10</sup>

sistema complejo de aprendizaje-enseñanza de lenguas, y ayudar a las instituciones que ofrecen cursos de lengua a ofrecer lo que hace falta, a nivel sociedad, institución e individuo.

Para cumplir esta función, el análisis prevé la recolección de datos por diferentes elementos del sistema, y acerca de diferentes áreas. La información a recolectar, se resume en el cuadro 2-1.

#### 2.1.6 El modelo de Munby

El modelo de Munby, presentado en su libro Communicative Syllabus Design<sup>11</sup> incluye la especificación más detallada que hemos visto de la forma que puede tomar un análisis de necesidades.<sup>12</sup> Un análisis de necesidades basado en el modelo de Munby parte del individuo, o, podemos interpretar, del grupo de participantes. Munby dice:

se empieza con la persona, y se investigan sus necesidades comunicativas particulares de acuerdo con las variables socioculturales y estilísticas que interactúan para determinar un perfil de tales necesidades.<sup>13</sup>

El modelo está diseñado para funcionar de manera lineal; se empieza con un participante, y se procesa información acerca de él y sus necesidades, para producir un 'perfil' del participante. Luego este perfil sirve como base para la especificación de necesidades comunicativas, en términos de micro-habilidades lingüísticas o de micro-funciones con realizaciones.

En la primera parte del modelo, en que se procesa información acerca del participante y se produce su perfil, Munby establece ocho parámetros:

- |                     |                        |
|---------------------|------------------------|
| 1. área de uso      | 5. dialecto            |
| 2. ambiente         | 6. nivel meta          |
| 3. interacción      | 7. evento comunicativo |
| 4. instrumentalidad | 8. tono comunicativo   |

Los parámetros uno a cuatro son parámetros a priori; dan información no lingüística necesaria para definir los elementos lingüísticos de los cuatro parámetros a posteriori (números cinco a ocho).

La información requerida para los ocho parámetros constituye el perfil de las necesidades comunicativas del participante. Basándose en este perfil, el que utiliza el modelo procede a establecer la competencia comunicativa meta. Aquí existen dos opciones: esta competencia meta puede definirse en términos de micro-habilidades, de micro-funciones con realizaciones específicas de ellas, o una combinación de las dos. Según Munby, la lista que resulta de la selección de estos elementos, con base en el perfil, es la definición del contenido del programa para los propósitos específicos del participante.

## 2.2 Discusión de las ventajas y desventajas de los modelos presentados

Para evaluar la utilidad de los diferentes elementos de los enfoques presentados en la parte 2.1, debemos considerar el contexto del trabajo actual. Aquí estamos considerando la situación hipotética de establecer un sistema de preparación

y evaluación para candidatos a becas en México. Las instituciones que otorgan becas, y que podrían interesarse en un sig tema de este tipo (CONACYT, Secretaría de Educación Pública, U.N.A.M.) son instituciones públicas. La consideración de la forma que debe tomar el análisis de necesidades se hace tenien do en cuenta esta situación global.

## 2.2.1 Los modelos de Strevens y de Munby

### 2.2.1.1 La taxonomía de tipos de cursos de lenguas para propósitos específicos

Un aspecto compartido de los dos modelos es la taxonomía de cursos de lengua para propósitos específicos que Munby tomó de Strevens, haciendo ciertas modificaciones. El uso de una taxonomía para establecer el tipo general de curso de que se trata es un elemento positivo en los dos modelos, ya que permite ubicar rápidamente el curso.

### 2.2.1.2 El estudio del contenido lingüístico

Un aspecto que incluye Strevens, el contenido lingüístico, re cibe muy poca atención en el trabajo de Munby, a nivel necesi dades. El establecimiento de realizaciones particulares a incluirse en el sílabo, el momento en que entra el contenido lingüístico en términos de léxico, sintaxis y especialmente características discursivas, ocurre en el modelo de Munby, después del establecimiento de micro-funciones y habilidades. Munby menciona la utilidad de estudios ya disponibles para es tablecer las realizaciones específicas a incluir (menciona los trabajos de Sinclair y Coulthard, y de Candlin, Leather y Bru-

ton, entre otros).<sup>14</sup> Sin embargo, no recomienda un estudio detallado del discurso apropiado, como parte del estudio de necesidades. En caso de que no haya descripciones ya existentes, recomienda tomar en cuenta las intuiciones, y las opiniones de informantes para llegar a decisiones acerca de qué contenido lingüístico incluir. Posteriormente, se puede refinar y modificar.

Mientras que sería ideal basar las decisiones acerca del contenido lingüístico de cursos específicos (en nuestro caso para becarios) en el estudio detallado de ejemplos de discurso del tipo que el participante tendrá que manejar, este tipo de análisis requiere de mucho tiempo, aún para analizar una cantidad mínima de texto. Una solución práctica a este problema es coleccionar textos que ejemplifican los tipos de texto relevantes para explotación en clases o exámenes, sin hacer análisis detallados de ellos. De esta manera, el estudiante está expuesto a las peculiaridades lingüísticas de los textos, y trabaja con ellas.<sup>15</sup>

### 2.2.1.3 La definición de objetivos en términos comunicativos

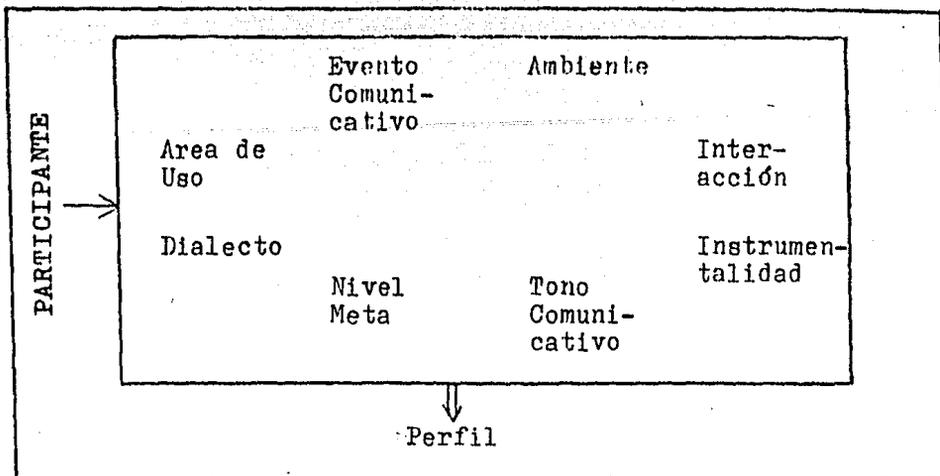
Un aspecto básico de los dos modelos es la definición de objetivos en términos comunicativos. Strevens usa el término 'categorías comunicativas'; Munby incluye micro-funciones y micro-habilidades. Las micro-habilidades forman una buena base para los objetivos académicos de los becarios, y se han utilizado para la especificación de objetivos de una parte del examen de requisito presentado en los capítulos IV y V.

#### 2.2.1.4 Problemas internos en el modelo de Munby

Nosotros adoptamos el modelo de Munby para especificar necesidades a nivel individual, por el gran detalle que ofrece en el establecimiento del perfil del participante, y por el inventario de micro-habilidades que presenta. Sin embargo, existen ciertas inconsistencias internas en el modelo, que hacen necesarios algunos cambios.

Munby trata de representar al proceso de diseño de cursos como un proceso lineal. Separa los ocho parámetros (véase la sección 2.1.6) en variables a priori y a posteriori, indicando una relación determinada, como si los parámetros a posteriori se derivaran de las variables a priori. Al aplicar el modelo, hay dificultad en establecer algunas de las variables a priori, sin conocer antes el evento comunicativo, según Munby, una variable a posteriori.

En general, sentimos que la linealidad propuesta en el modelo de Munby no refleja la realidad. De hecho las diferentes variables influyen la una en la otra de manera compleja y no determinada. Por lo tanto, presentamos un modelo modificado del procesador de necesidades comunicativas de Munby, que simplemente presenta los factores que intervienen, sin especificar relaciones determinadas entre ellos (cuadro 2-2).



CUADRO 2-2

MODIFICACION DEL PROCESADOR  
DE NECESIDADES COMUNICATIVAS DEL MODELO DE MUNBY

2.2.1.5 Otros niveles

El modelo de Munby (que contiene, y elabora varios aspectos del modelo de Strevens) funciona, con modificaciones, para establecer necesidades al nivel individuo. Sin embargo, existen otros niveles en los cuales debe realizarse un análisis de necesidades; hay otros factores a considerar, aparte del participante.

Munby en su epílogo (que consta sólo de dos páginas) menciona una serie de factores que él admite entran en el proceso del diseño de un programa.<sup>16</sup> Son los aspectos sociopolítico, práctico, administrativo-educativo, psicopedagógico, y metodológico. Sólo aquí incluye Munby esta serie de factores tan importantes en el establecimiento de un curso en el mundo real. Haberlos excluido de su modelo, considerándolos como posteriores al diseño del programa parece inconsistente en un

libro cuyo subtítulo indica que da un "modelo sociolingüístico" para el diseño de cursos. Un modelo verdaderamente sociolingüístico no percibe el problema de la definición de necesidades exclusivamente en el nivel del individuo, o aún del grupo de participantes. Otros modelos que examinamos arriba, los de MacKay, Richterich y Chancerel, y Shaw, por ejemplo, toman un acercamiento al análisis de necesidades que incluye otros factores muy importantes. Elementos de estos enfoques deben complementar el modelo de Munby para adecuarlo al análisis de necesidades para becarios mexicanos.

### 2.2.2 El Consejo de Europa

El modelo del Consejo de Europa tiene algunas características específicas que vienen del hecho de que se diseñó para usarse en un sistema completo regional de enseñanza-aprendizaje de lenguas, y que por lo tanto, no son aplicables al presente análisis. Sin embargo otros aspectos, como la consideración del análisis de necesidades como proceso continuo son generalizables.

### 2.2.3 Contribuciones de MacKay, Brumfit y Shaw

Las observaciones de MacKay, Brumfit y Shaw contribuyeron al modelo utilizado en este trabajo en el sentido que todos de una manera u otra abogan por un análisis que tome en cuenta no solamente las necesidades individuales, sino toda una gama de factores que influyen en el diseño de cursos e instrumentos de evaluación.

### 2.3 El modelo de análisis de necesidades utilizado en el presente trabajo

El modelo que hemos adoptado para este análisis es uno que incluye varios de los factores tomados en cuenta en los modelos revisados en la sección 2.2.

El primer paso del análisis de necesidades consiste en la delimitación de los objetivos globales, y la política general, con las necesidades institucionales. Este estudio corresponde a la justificación del curso. Aquí se demuestra que social o institucionalmente, existe la necesidad de que haya conocimientos del inglés del tipo que se intenta dar. Este proceso es análogo a la investigación hecha por MacKay, para determinar la necesidad institucional de cursos de comprensión de lectura en veterinaria.

Luego de encontrar que el curso estaría justificado de esta manera, procedemos al segundo paso - la evaluación de la situación actual para asegurar que el curso no será redundante. Esto implica una evaluación del estado de conocimientos de la población meta: si ya están preparados en el área, no es necesario el curso. También, aquí se analizan los cursos ya existentes para ver si no existe ya un curso adecuado para la población

Finalmente, ya establecida la necesidad del curso en estos niveles, se procede a delimitar con exactitud las necesidades,

comunicativas del participante, a partir de un perfil del estudiante. Aquí, el modelo de Munby, con ciertas modificaciones, sirve de base.

Nuestro enfoque al análisis de necesidades entonces, tiene las siguientes características: se basa en una justificación sociopolítica e institucional; contiene un elemento de evaluación de conocimientos de la población meta, y los cursos ya existentes; y llega al nivel de análisis de necesidades comunicativas del participante sólo después de haber establecido en otros niveles, la necesidad del curso.

---

#### Referencias bibliográficas

#### Capítulo II

<sup>1</sup>Peter Strevens, New Orientations in the Teaching of English (Oxford: Oxford University Press, 1977), pp. 89-108.

<sup>2</sup>Ibid., p. 93.

<sup>3</sup>Ibid., p. 107.

<sup>4</sup>Ibid.

<sup>5</sup>Ronald MacKay, "Identifying the Nature of the Learner's Needs," en English for specific Purposes, eds. Ronald MacKay y Alan Mountford (London: Longman, 1978), pp. 21-42.

<sup>6</sup>Christopher Brumfit, "'Communicative' Language Teaching: An Educational Perspective," en The Communicative Approach to Language Teaching, eds. C. J. Brumfit y K. Johnson (Oxford: Oxford University Press, 1979) p. 185.

<sup>7</sup>Anthony Shaw, "1.2/1 Diseño de Cursos" Manuscrito, Biblioteca del CELE, UNAM., p. 6.

<sup>8</sup>René Richterich y Jean-Louis Chancerel, Identifying the Needs of Adults Learning a Foreign Language. (Oxford: Pergamon Press, Council of Europe, 1977).

<sup>9</sup>Ibid., p. 6

<sup>10</sup>Ibid., pp. 7-8.

<sup>11</sup>John Munby, Communicative Syllabus Design: A Socio linguistic model for defining the content of purpose specific language programmes (Cambridge: Cambridge University Press, 1978).

<sup>12</sup>Vid. Gwenn Colle, Ann Hildreth y David Howard, "Tra bajo Semestral para Diseño de Cursos (CELE/ UNAM, 1981), y Ann Hildreth y David Howard, Reseña de Communicative Syllabus Design de John Munby en Estudios de Lingüística Aplicada, en prensa.

<sup>13</sup>Munby, Communicative Syllabus Design,

<sup>14</sup>Ibid, pp. 49-50.

<sup>15</sup>Anthony Shaw, personal communication

<sup>16</sup>Munby, Communicative Syllabus Design, pp. 217-218.

## CAPITULO III

### LOS RESULTADOS DEL ESTUDIO DE NECESIDADES

#### 3.0 Introducción

Aquí se resumen los resultados del estudio de necesidades que forma la base del instrumento de evaluación propuesto en este trabajo. La mayor parte de la información aquí presentada es de un estudio realizado y reportado en 1981.<sup>1</sup>

#### 3.1 Justificación - necesidades nacionales e institucionales

El gobierno de México, a través sobre todo del CONACYT realiza un programa de desarrollo que tiene como meta la autodeterminación científica y tecnológica. Este programa incluye el otorgamiento de becas para estudiar en el extranjero como mecanismo para llegar a esta meta. En los últimos años (hasta 1982), el crecimiento en el número de becas había sido continuo. La situación económica de México a partir del verano de 1982 afecta los programas de becas, disminuyendo su número. Sin embargo, sigue habiendo algunas becas, probablemente en áreas más cuidadosamente escogidas como esenciales. Consideramos que el número reducido de becas no implica que un sistema de preparación y evaluación para los candidatos deje de tener importancia. De hecho, con menos recursos, es aún más

importante asegurar que los becarios que van al extranjero, vayan listos para enfrentar las dificultades de hacer estudios en inglés. Así se trataría de evitar que las personas pasen varios meses o hasta un año en el extranjero, con la actividad principal de aprender el idioma.

Mientras que se ofrecen becas para hacer estudios en un gran número de países, la mayoría (75 % de los becarios de CONACYT<sup>2</sup> y 67 % de los de la UNAM<sup>3</sup>) que estudian en países donde no se habla español, lo hacen en países de habla inglesa.

La importancia de los estudios en el extranjero para el desarrollo de México, en la opinión de investigadores y científicos sobresalientes fue comprobada en entrevistas personales de Howard con varios de ellos.<sup>4</sup>

Estos datos nos parecen demostrar a niveles nacional e institucional, la necesidad que hay que los candidatos a becas tengan dominio del inglés suficiente para realizar estudios de posgrado en países de habla inglesa. Esto constituye la justificación básica del sistema que prevemos.

## 3.2 Evaluación

### 3.2.1 Evaluación del nivel de la población meta

Hay varias indicaciones de que los candidatos a becas actualmente no tienen el dominio necesario para llevar a cabo sus estudios en el extranjero.

El individuo que quiere recibir una beca para estudiar en el extranjero debe presentar al organismo que otorga la beca un comprobante de que ha sido aceptado en una universidad. Esto en principio, implicaría que ya hubiera cubierto ciertos requisitos mínimos en cuanto al idioma, ya que la mayoría de las universidades de Norteamérica y Gran Bretaña requieren exámenes de inglés, o sea el TOEFL o el examen del "English Testing Service." Sin embargo, tener la aceptación de una universidad no garantiza que los estudiantes pueden aprovechar sus cursos. El valor del examen TOEFL se ha puesto en duda, y el director de un centro de lenguas de una universidad de los Estados Unidos, encargado de preparar estudiantes extranjeros para las exigencias del estudio profesional y de posgrado en inglés ha dicho que "el TOEFL no significa nada."<sup>5</sup> Y en Gran Bretaña, la existencia de un gran número de cursos propedeúticos indica que las autoridades de las universidades están concientes del hecho de que están aceptando a estudiantes que necesitan más entrenamiento antes de entrar a sus programas académicos. CONACYT tiene programas establecidos con dos universidades en los Estados Unidos de Norteamérica por medio de los cuales sus becarios reciben cursos de lengua en los Estados Unidos. En uno de estos programas, los becarios en economía pasan hasta nueve meses en Boulder, Colorado, antes de que lleguen siquiera a la universidad en que van a hacer sus estudios de posgrado.

Además, la información recolectada por medio de entrevistas con ex-becarios, y los comentarios de maestros en diferentes instituciones que preparan becarios señalan que generalmente existen deficiencias serias, aún en personas que las universidades en el extranjero ya han aceptado.

Estos hechos nos llevan a la conclusión que sí existe una necesidad de un sistema de preparación y evaluación.

### 3.2.2 Información acerca de cursos ya existentes

Aunque existen algunos cursos para personas que se quieren preparar para hacer estudios en países de habla inglesa, no son suficientes, y no están diseñados especialmente para ayudar al estudiante a alcanzar el nivel que implica la tarea que enfrentará.

En la UNAM existe el examen de requisito para becarios, que pretende garantizar que los becarios tengan un dominio suficiente de la lengua. Sin embargo, no hay cursos correspondientes para los candidatos que no tengan el nivel requerido. Los candidatos deben acomodarse en grupos de inglés donde el nivel general está debajo de lo que se pide en el examen, y donde el programa no prevé específicamente las necesidades de ellos.

CONACYT tiene un centro de lenguas para preparar a los candidatos a becas de esta institución, en las diversas lenguas.

implicadas en el programa de becas. Desafortunadamente, al menos en inglés, problemas de espacio, personal y presupuesto hacen que sea imposible que el centro dé un servicio que incluya un largo período de estudio para cada candidato. Problemas de cupo hacen que a veces los candidatos que necesitan empezar sus estudios a nivel principiante, deben esperar uno o más bimestres para comenzar, porque no hay lugar para él. Esto afecta el nivel que podrá alcanzar el candidato, ya que tiene un período limitado antes de salir al extranjero para empezar su programa académico. Así es que los cursos de lengua en CONACYT no pueden satisfacer del todo las mismas necesidades de sus propios becarios, mucho menos tratar de ayudar a personas de otras instituciones.

En el caso de CONACYT el problema de cupo y falta de tiempo, se relaciona con otro problema, la naturaleza misma de los cursos. Debido al bajo nivel de los candidatos al momento de ingresar al centro de lenguas, y el tiempo reducido que tiene el becario antes de salir, los cursos básicos se orientan hacia objetivos de sobrevivencia en el extranjero. Sólo en los cursos avanzados (a donde sólo una minoría de los candidatos llegan), se trabaja sobre las habilidades de estudio que forman gran parte de las necesidades específicas de estos candidatos. Aunque los cursos básicos son de gran importancia para CONACYT, no pueden considerarse como una solución a los problemas de lengua de los becarios.

En la Universidad Ibero-Americana, se dan cursos para preparar a personas que van a tomar el examen TOEFL. Estos cursos preparan al alumno para contestar a preguntas de opción múltiple, como los que aparecen en el examen, y no tienen como objetivo preparar al alumno a enfrentarse a las situaciones comunicativas que encontrará cuando está estudiando en el extranjero. Hace varios años, esta misma universidad ofrecía un curso "Pre-Masters" diseñado para ayudar al alumno a adquirir habilidades de estudio. Para entrar en el curso, los estudiantes tenían que haber recibido ya una calificación aceptable para universidades norteamericanas en el examen TOEFL. Es interesante notar que el curso se dejó de dar por falta de interés en los estudiantes. Aparentemente, ya aceptados en la universidad en el extranjero, estos estudiantes creyeron que su inglés era suficiente para hacer sus estudios. Como hemos visto, esta opinión parece ser errónea. Parece que para implantar un sistema de preparación que siga más allá que la aceptación en la universidad extranjera, y llevar al candidato a un dominio superior al mínimo requerido por la universidad, será necesario que las instituciones que otorgan becas pongan requisitos superiores a este mínimo. La experiencia de la UIA indica que un programa voluntario puede fracasar por falta de interés de los estudiantes.

Se ve que dado el nivel insuficiente en lengua que frecuentemente presentan candidatos a becas, y la falta general de formas institucionales para preparar candidatos, y evaluar sus

conocimientos sí sería justificada la implantación de sistemas de preparación y evaluación de becarios.

### 3.3 El perfil del participante

Participante(s): Hombres y mujeres mexicanos, de 25 a 35 años de edad, aproximadamente. Durante el curso de idioma, y en el momento del examen vivirán en la República Mexicana, probablemente en el Distrito Federal.

Area de Uso: Inglés para propósitos académicos, específicamente para estudiar programas de posgrado en una variedad de disciplinas académicas.\*

### Eventos y Actividades Comunicativos

Fase I - Comprensión auditiva (conferencias)\*\*

Evento 1 - Asistir a conferencias semi-informales e informales en inglés

Actividad 1.1 Escuchar para la comprensión de las ideas principales.

Actividad 1.2 Escuchar para información específica.

---

\* Para concentrar más los esfuerzos en este trabajo, se ha decidido trabajar exclusivamente con las necesidades que distinguen al becario de otros tipos de participante; sus necesidades en el área de uso académico. Sus necesidades de sobrevivencia (interacciones sociales, comerciales, etc.) las comparte el becario parcial o totalmente con otros tipos de participantes.

\*\* La organización de los eventos y las actividades en cuatro fases fue posterior al análisis de necesidades, y refleja decisiones acerca de la organización de la evaluación.

Actividad 1.3 Escuchar para saber cuál es la actitud del conferencista hacia el tema.

Actividad 1.4 Tomar apuntes para referencia en el futuro.

Actividad 1.5 Leer material escrito en el pizarrón o en hojas distribuidas antes de, o durante la conferencia.

Actividad 1.6 Pedir clarificación.

## Fase II - Comprensión de lectura

Evento 2 Estudiar materiales de referencia (asignados) en inglés; libros de textos, trabajos monográficos y artículos de revistas especializadas.

Actividad 2.1 Leer para entender la mayoría de la información en el texto.

Actividad 2.2 Leer, sacando sólo los puntos principales.

Actividad 2.3 Leer para saber la actitud del autor hacia el tema.

Actividad 2.4 Tomar apuntes para referencia en el futuro.

Actividad 2.5 Subrayar partes claves de un texto para facilitar referencia en el futuro.

Evento 3 Consultar material nuevo relevante (no asignado)

Actividad 3.1 Revisar material para decidir su relación con el área para evaluar su importancia.

Evento 4 Hacer investigación bibliográfica.

Actividad 4.1 Consultar libros de referencia.

Actividad 4.2 Utilizar aparatos para lectura de micro-formas.

Actividad 4.3 Utilizar fichero.

Actividad 4.4 Consultar con el bibliotecario de referencia.

### Fase III - Producción escrita

Evento 5 Escribir reportes, trabajos y tesis en inglés (incluye el evento 4, hacer investigaciones bibliográficas).

Actividad 5.1 Escribir el borrador.

Actividad 5.2 Hacer correcciones en el borrador.

Evento 6 Tomar exámenes escritos.

Actividad 6.1 Estudiar para el examen: revisar apuntes, lectura, etc., sacando conceptos y puntos claves.

Actividad 6.2 Organizar información e ideas bajo presiones de tiempo

Actividad 6.3 Redactar bajo presiones de tiempo.

### Fase IV - Interacción oral (seminario)

Evento 7 Participar en seminarios y pláticas de grupo.

Actividad 7.1 Organizar y presentar información y opiniones sobre un tema preparado con anterioridad.

Actividad 7.2 Aclarar dudas, reformulando, dando ejemplos, etc.

Actividad 7.3 Argumentar, defender un punto de vista, acceder a otro punto de vista en su caso.

Actividad 7.4 Participar en una discusión basada en presentaciones hechas por compañeros.

Ambiente: Ambiente físico: Participantes harán sus estudios principalmente en los Estados Unidos de Norteamérica, Gran Bretaña y Canadá. En la universidad, académicamente hablando usará el inglés en el salón de conferencias, el salón de clases, el laboratorio, el seminario, los cubículos de sus

profesores y tutores, en la biblioteca y en su casa. Ambiente psico-social. El ambiente universitario en el extranjero es parecido culturalmente, al ambiente universitario que conoce aquí en México. Es un ambiente intelectual, ético, altamente desarrollado educacional y tecnológicamente, urbano, aparentemente tranquilo, pero con grandes presiones en el trabajo académico. Dependiendo de características particulares del departamento en que trabaje, encontrará relaciones formales jerárquicas, y autoritarismo, o relaciones informales e igualitarias. Ambiente físico-temporal. Los participantes usarán el inglés constantemente; el uso académico de esta lengua probablemente ocupará por lo menos cuarenta horas por semana.

Interacción: Como estudiante de posgrado, interactuará con hombres y mujeres, (con nativos del país donde estudia y con otros extranjeros), que serán sus profesores y compañeros. Esta interacción implicará una serie de relaciones complejas en que su compañero es a la vez amigo íntimo y competidor; en que un profesor puede considerarlo como miembro de su "grupo" y otro lo considera como del exterior.

Instrumentalidad: Los participantes tendrán que entender la palabra hablada y escrita, y tendrán que expresarse oralmente y por escrito en situaciones de monólogo como conferencias, libros y trabajos, y diálogos como seminarios, pláticas y correspondencia. Manejarán la interacción cara-a-cara, por

teléfono e impresa, dominando aspectos tanto verbales como no verbales.

Dialecto: El participante tendrá que entender, y producir de manera inteligible el inglés estándar contemporáneo, y entender una variedad de acentos locales y no-nativos del inglés.

Nivel meta: Para una consideración del concepto de nivel meta, y su especificación para el examen de requisito, véanse las secciones 5.4.2 y 5.5.

Tono comunicativo: Durante sus estudios de posgrado, el estudiante extranjero se encontrará con toda una serie de actitudes, y por lo tanto, de tonos comunicativos. A lo largo de sus estudios, podrá ir entendiendo sutileza cada vez mayor. Básicamente será necesario que pueda distinguir entre una actitud positiva y una actitud negativa, ante el tema a tratar del hablante/escritor. Para hacer esto, a nivel receptivo, deberá distinguir entre: admirar/despreciar y aprobar/desaprobar. A nivel productivo, para participar mínimamente en una discusión tendrá que mostrar su acuerdo o desacuerdo, y su seguridad o inseguridad en cuanto a lo que otro ha dicho, o en cuanto a su contribución. Para esto a nivel receptivo y productivo, tendrá que distinguir entre: seguridad/inseguridad y acuerdo/desacuerdo.

## Referencias bibliográficas

### Capítulo III

<sup>1</sup>Gwenn Colle, Ann Hildreth y David Howard, Proyecto Semestral para Diseño de Cursos (CELE/UNAM, 1981).

<sup>2</sup>Ibid.

<sup>3</sup>Universidad Nacional Autónoma de México, Boletín de Información para Becarios en el Extranjero. I,2 (Marzo, 1982).

<sup>4</sup>Colle, et al.

<sup>5</sup>Entrevista con el Dr. Robert Fox, American University, Washington, D. C.

## CAPITULO IV

### DEFINICION Y DISCUSION DE CONCEPTOS

#### BASICOS EN LA EVALUACION

##### 4.0 Introducción

La evaluación inevitablemente encierra una serie de problemas que no tienen solución, por lo menos a corto y mediano plazos, problemas que harían que uno dejara la consideración de la evaluación hasta un punto en el futuro indefinido, diciendo "todavía falta información para comprobar la validez del examen" y "no es justo basar decisiones en el instrumento que tenemos, aunque sea el mejor posible en este momento." Este suele ser el caso, más con la evaluación que con otros aspectos del proceso educativo, porque es imposible ser tentativo en las decisiones que se hacen en base a instrumentos de evaluación, y los resultados de la evaluación son más tangibles y visibles que los resultados de otros aspectos, como son el diseño del sílabo y la metodología. Es más fácil llegar a la conclusión que la metodología que usamos en un momento dado no es la adecuada, que no está dando los resultados esperados, y hacer cambios, que confesar que un examen no mide lo que pretendíamos medir, que no han aprobado personas que deberían de aprobar y vice versa. En el caso de la metodología, el problema puede quedar "entre nosotros", pero en

el segundo caso, podemos recibir reclamaciones por parte de personas quienes han reprobado el examen que ahora no consideramos adecuado.

Problemas de este tipo hacen de la evaluación un tema sumamente delicado, y difícil de abordar. Sin embargo, otros factores la hacen una de las áreas más esenciales del proceso educativo, ya que la evaluación es el eslabón, nos guste o no, entre un sistema educativo, y otros sistemas.

En la situación de la preparación de becarios para estudios en países de habla inglesa, el "sistema" completo para el otorgamiento de becas requiere de información sobre candidates. Puede suceder (como es el caso en la U.N.A.M.) que este sistema de otorgamiento de becas tenga un requisito institucional de cumplimiento con un nivel mínimo de conocimientos de lengua para poder recibir una beca para hacer estudios en el extranjero. En este caso se necesita una evaluación de los conocimientos del idioma del candidato, para el uso de los que van a tomar decisiones acerca del otorgamiento o no de una beca.

Para diseñar un examen de requisito para esta situación, se considera el propósito del examen: evaluar el grado en que el candidato está preparado lingüísticamente para enfrentarse a las exigencias de un programa académico de posgrado en un país de habla inglesa. Intentar establecer esto por medio

de un examen es una tarea muy difícil - tal vez imposible. Una gama de factores, algunos lingüísticos, otros no lingüísticos, llegarán a influir en el éxito o falta de éxito del becario en el extranjero. La soledad, y la dificultad para establecer relaciones sociales - cosas que pueden tener más relación con la personalidad del becario, y el medio en que se encuentra, que con dificultades con la lengua - pueden contribuir al no aprovechamiento de una beca. Sin embargo, a la vez, los estudiantes extranjeros frecuentemente compensan las faltas de conocimiento de la lengua, dedicando más horas a la lectura, y utilizando otras estrategias para salir adelante. Estas dificultades nos podrían llevar a la conclusión que no es posible predecir el éxito de un individuo en un programa de estudios en el extranjero en base a un examen de idioma. Sin embargo, hace falta intentar hacerlo. A pesar de los problemas inherentes en la evaluación de candidatos a becas, es necesario tratar de encontrar un instrumento que nos dé información precisa acerca del manejo de idioma del candidato, en relación a la situación académica de programas de posgrado en el extranjero.

En este capítulo se discuten algunos conceptos básicos en la evaluación, definiendo algunos términos, y se revisan algunos estudios en el área, comentando cuando se relacionan con decisiones específicas en cuanto a la forma que toma el esquema del examen de requisito para becarios.

#### 4.1 Tipos de reactivos y exámenes

##### 4.1.1 Reactivos de puntos aislados versus exámenes integra- tivos

El reactivo de punto aislado está basado en un análisis lingüístico de la lengua, en las áreas de fonología, morfología, sintáxis y semántica, y mide el dominio de un elemento específico de la lengua meta. Poderlo contestar no debe depender de la comprensión de un contexto amplio, sino sólo del punto específico a evaluar. Este término, entonces, se refiere no a la forma que toma el reactivo, sino al contenido que se está evaluando con ello. Unos ejemplos pueden aclarar la cuestión.

Davies describe un examen que evalúa si el examinando domina a nivel productivo ciertas distinciones fonéticas, el "McCa-  
llien Oral English Test for West Africa".<sup>1</sup> El examinando debe leer un texto en voz alta; el texto se ha preparado de manera que contenga las distinciones fonéticas a examinar. El examinador sólo evalúa la producción de estas distinciones, y ninguna otra cosa; fluidez, rapidez, pronunciación de otros fonemas, entonación, ritmo, etc. quedan fuera de la evaluación. Este examen es de puntos aislados. Obviamente, aquí, la evaluación no puede llevarse a cabo automáticamente; el examinador debe entender completamente los puntos fonéticos a evaluar, y tal vez emitir un juicio un tanto subjetivo acerca de la realización de la distinción fonética a evaluar. Mientras los fonemas son discretos, su producción por hablantes

no nativos puede cubrir un continuo que va desde una realización completamente equívoca hacia una realización parecida a la de un hablante nativo. El examinador debe decidir si la realización de un examinando está suficientemente cercana a la ideal para considerarse como aceptable. Pero es un ejemplo de un examen de puntos aislados, ya que es claro cuál elemento lingüístico se está evaluando, y se eliminan, a la hora de la calificación, dificultades que no se relacionan con esto.

Los reactivos en un examen de puntos aislados se presentan frecuentemente sin contexto, pero esto no es su característica definitiva. Pueden tener un contexto, pero el hecho de contestar no depende de la comprensión del contexto amplio. Citamos un ejemplo de Valette.

(1) Situational context. Richard is out of luck! Whenever he asks his friends to do something, he finds out they have already done so. Complete the answers with the past tense form of the underlined verb.

1. Richard: Do you want to go to the movies?  
Sally: Sorry, I \_\_\_\_\_ to the movies 2  
last night.

Este reactivo "contextualizado" evalúa un punto aislado - la forma pasado del verbo "go". Requiere del examinando lo mismo que un reactivo de este tipo:

(2) Write the past tense form of these verbs:

1. go \_\_\_\_\_  
2. see \_\_\_\_\_

Sólo se evalúa un elemento morfológico en estos dos ejemplos, aunque en el caso del reactivo de Valette (1), también entra la capacidad del alumno de comprender las instrucciones y no



Realmente es imposible definir exactamente cuando un reactivo deja de ser de punto aislado. Casi todo reactivo requiere de la comprensión de un contexto mínimo. Reactivos que evalúan conocimientos acerca de tiempos verbales generalmente requieren que el examinando relacione expresiones como "yesterday" y "last week" con el pasado simple, o adverbios de frecuencia con el presente simple, aunque se les puede considerar como reactivos de puntos aislados.

Oller parece distinguir entre una teoría de puntos aislados, que dice que estos reactivos deben medir un solo punto de la lengua a la vez, y la realidad de los reactivos posibles. En la realidad, dice Oller, existe "un continuo entre los reactivos de punto aislado en un extremo, y el uso total de la lengua en el otro."<sup>3</sup> Tratar de trazar el límite del reactivo de punto aislado sería un ejercicio sin sentido real. Lo importante al evaluar cualquier reactivo es analizarlo con cuidado para determinar exactamente qué requiere del examinando en términos de comprensión, conocimiento y producción. Con esta información, el examinador podrá decidir si el reactivo es adecuado en base a sus objetivos en el examen.

Oller define el examen integrativo como uno que, en contraste con el examen de puntos aislados, evalúa la capacidad del examinando de utilizar varios elementos a la vez.<sup>4</sup> Puede ser que esto implique el uso de más de una habilidad, es decir puede requerir que el examinando lea y escriba, o que escuche

y hable. Preguntas de comprensión global (que no dependen de la comprensión de un elemento léxico, fonético, morfológico, o sintáctico particular, sino del contexto completo) son integrativos, como son el dictado y el examen cloze. Es de notar que el examen integrativo no necesariamente refleja una tarea normal o uso comunicativo de la lengua. Sin embargo, en el continuo entre reactivos de puntos aislados, y uso total, la prueba integrativa se acerca un poco más al lado de uso.

En su tratamiento de exámenes integrativos, Oller introduce otra característica que puede tener un examen: puede ser pragmático. Un examen pragmático evalúa, según Oller, lo que él llama la "gramática de expectativas" del examinando - su capacidad de manejar la lengua con las restricciones de tiempo que rigen el uso normal de la lengua, y relacionarla con su contexto extralingüístico. Trata de medir el grado en que el examinando tiene elementos suficientes de la lengua para predecir, planear y evaluar lo que está recibiendo o produciendo, y adecuarlo al contexto. Para evaluar estos elementos, el examen pragmático debe 1) estar sujeto a restricciones de contexto y de tiempo y 2) requerir que el examinando relacione secuencias lingüísticas a contextos extralingüísticos de manera significativa.<sup>5</sup> Mientras que todo examen pragmático será integrativo (ya que debe depender de un contexto amplio), no todo examen integrativo será pragmático. Sin el factor de restricción de tiempo, un examen podrá ser integra-

tivo, mas no pragmático.

#### 4.1.2 Reactivos o exámenes objetivos y subjetivos

Diferentes autores toman el término "objetivo" en diferentes sentidos. Para Schultz, un reactivo objetivo parece ser lo mismo que un reactivo de punto aislado. Dice:

Un examen de puntos aislados (generalmente llamado "reactivo objetivo") se define como cualquier reactivo independiente, no limitado por elementos situacionales, que requiere de una respuesta específica, predeterminada.

Davies le da otro uso al término cuando dice:

Una prueba estandarizada es objetiva: esto quiere decir que ... es tanto confiable como válida.<sup>7</sup>

Aquí se rechazan estos usos del término "objetivo" en favor de la definición que da Pilliner. Según esta definición, el hecho de que un reactivo sea objetivo o no, depende sólo del aspecto de la evaluación de la respuesta del examinando. La calificación de un reactivo se considera objetiva siempre y cuando la decisión acerca de la adecuación de la respuesta sea automática, y cada respuesta no requiera de un juicio deliberado.<sup>8</sup> Como utilizamos el término, entonces, la objetividad de un reactivo o un examen se refiere sólo a la calificación de la respuesta como adecuada o no, y no a la tarca que se pide del examinando, ni la naturaleza de lo que se está evaluando con el reactivo, ni a su calidad en términos de su buen funcionamiento como instrumento de evaluación.

El reactivo o examen objetivo, entonces, puede tomar diferentes

formas, y puede evaluar diferentes cosas; seguirá siendo objetivo mientras la calificación de cada examen no requiera de una decisión meditada. Claramente, cualquier reactivo de opción múltiple o falso-verdadero es objetivo, a pesar del hecho de que reactivos de este tipo pueden medir tanto puntos aislados como tareas muy complejas que requieren el manejo de un contexto amplio. Una pregunta de comprensión de lectura que pide al examinando que decida cuál, de cuatro oraciones, es el mejor resumen del punto principal de un texto largo, es objetivo, ya que su calificación es automática, aunque evalúa un complejo de habilidades y conocimientos lingüísticos. Un examen de tipo cloze también es objetivo, según esta definición. Basta señalar que existen programas de computadora para la calificación de exámenes de este tipo.<sup>9</sup>

A diferencia del examen objetivo, el examen subjetivo sí implica la aplicación de criterios que requieren una consideración meditada y un juicio sobre la actuación de cada individuo que toma el examen.

Hay ventajas y desventajas en ambos tipos de exámenes (objetivos y subjetivos). La calificación objetiva, naturalmente, se relaciona con un alto grado de confiabilidad\* en la corrección del examen. Difícilmente habrán diferencias entre las evalua-

---

\* Véanse las secciones 4.3 y 4.4 para definiciones de los términos confiabilidad y validez.

ciones de dos examinadores, calificando un examen objetivo. Con exámenes subjetivos, hay la posibilidad de que dos examinadores evalúen la misma actuación de manera distinta, o que el mismo examinador evalúe la misma actuación de dos maneras distintas, en diferentes momentos, hecho que afecta la confiabilidad de la medida. Otra ventaja del examen objetivo es la rapidez con la que se califica. El examen subjetivo, que requiere juicios meditados, naturalmente requiere más tiempo para su calificación. También necesita un examinador altamente preparado para la tarea, mientras que cualquier persona que tiene una clave que le indique cuáles son las respuestas correctas puede calificar una prueba objetiva.

La prueba objetiva por otra parte, tiene algunas desventajas. Para poder calificar una prueba de manera objetiva, debe haber una sólo respuesta correcta, o un número finito, y preferiblemente bastante reducido, de respuestas posibles. Por esta razón generalmente la prueba objetiva es en forma de opción múltiple o falso-verdadero. Mientras este tipo de prueba puede ser útil para medir comprensión y reconocimiento, no es posible medir la producción (ni oral ni escrita) de manera directa, utilizando un formato objetivo. Otra desventaja de la prueba objetiva es que puede ser de baja confiabilidad en cuanto a la re-examinación del individuo. Debido al hecho de que el examinando podrá adivinar cuando no sabe cuál es la respuesta correcta, y a veces adivinar correctamente, el resultado del mismo individuo puede cambiar entre una y otra administración del examen, aunque sus conocimientos no hayan

cambiado. Esto demuestra que los resultados del examen objetivo no necesariamente reflejan los conocimientos del individuo examinado.

El peso de la tradición en la evaluación, que da un valor tal vez exagerado a la confiabilidad y economía del examen, ha hecho que los examinadores frecuentemente insistan en formatos que permitan la calificación objetiva, aunque ésto signifique la exclusión de sub-pruebas para medir producción oral y escrita. En estas situaciones, consideraciones de practicabilidad y confiabilidad han superado consideraciones acerca del objetivo del examen, haciendo que la validez del examen se reduzca en gran medida.

Decisiones de este tipo frecuentemente se justifican, citando la regla básica del análisis estadístico de exámenes, "el examen no puede ser válido si no es confiable". Este hecho puede llevar al diseñador de un examen a pensar antes que nada en la confiabilidad de su instrumento, y a optar por tareas que se califiquen objetivamente, para lograr la alta confiabilidad en la calificación, dejando fuera otras tareas en que este tipo de confiabilidad es más difícil de lograr.

En el diseño del examen de requisito, se preve la inclusión de formatos de calificación objetiva por las ventajas antes mencionadas (economía y confiabilidad de calificación), para la evaluación de ciertos tipos de conocimientos.

Sin embargo, también se incluyen partes cuya calificación es subjetiva (sobre todo en la evaluación del nivel comunicativo,

en el lado productivo). Aunque se reconoce la importancia de la confiabilidad en la calificación, permitimos que la validez de la medida (el grado en que mide lo que pretende medir) tomara importancia primaria en el diseño original. De otra manera, correríamos el peligro de obtener un examen confiable pero que no proporcionara información que se necesita acerca del candidato.

#### 4.1.3 Exámenes de aprovechamiento

El término examen de aprovechamiento se refiere al uso que se le da a la prueba. El examen de aprovechamiento se utiliza para medir el grado en que los alumnos han logrado los objetivos de un curso específico. El contenido, entonces, es un reflejo del contenido del curso, y como tal podrá ser de puntos aislados, integrativo, o comunicativo dependiendo de lo que se quiso lograr con el curso. Los resultados del examen de aprovechamiento frecuentemente se usan para decidir si el examinando "pasa" el curso. También la información arrojada por este tipo de examen puede informar acerca de la efectividad de los métodos de enseñanza, e influir en el curso, en el futuro.

#### 4.1.4 Examen de requisito

Aquí se entiende por examen de requisito, aquel examen que se

le pide a una persona como requisito para realizar una actividad que requiere el uso de la lengua. Un ejemplo sería un examen de inglés que hace una compañía a candidatos para empleos; los candidatos deben pasar el examen para ser considerados para el trabajo. Supuestamente, el propósito del examen es asegurar que el candidato tiene suficiente dominio del idioma para llevar a cabo ciertas tareas en el trabajo que requieren del uso del inglés. Claro está que el examen que se hace a candidatos a becas para decidir si están preparados lingüísticamente para hacer estudios en el extranjero cabe en esta definición.

El examen de requisito, entonces, es un examen utilizado para determinar si el candidato podrá usar la lengua para sus necesidades futuras. Notamos que este término se refiere al uso administrativo del examen, y no al examen en sí. Es posible, y desafortunadamente, frecuente, que se use un examen como requisito, es decir para juzgar la habilidad de la persona en cuanto a su posibilidad de llenar sus futuras necesidades en el idioma, siendo el examen en sí un instrumento sumamente inadecuado para este propósito.

#### 4.1.5 Examen de dominio

El examen de dominio pretende medir un factor "L", - conocimientos de la lengua, fuera de cualquier curso de idiomas y sin hacer referencia a ninguna situación particular comunicativa. Puede tomar diferentes formas, por ejemplo puntos aislados o integrativo, y puede constituirse de una sola prueba

(cloze o dictado, por ejemplo), o puede contener una serie de sub-pruebas para diferentes habilidades. En todo caso, básicamente mide el conocimiento que tiene el examinando de la lengua. Sin embargo, los exámenes de dominio generalmente tienen un propósito más amplio. Pretenden dar información que va más allá que el contenido directamente medido. En palabras de Davies,

el dominio de un idioma implica un control suficiente de habilidades lingüísticas para un propósito extralingüístico.<sup>10</sup>

La implicación es que los resultados de un examen de dominio darán información acerca de cosas que no mide directamente, acerca de la habilidad del examinando para usar su dominio de la lengua, en situaciones comunicativas, o sea, para propósitos extralingüísticos. Se hace entonces una extrapolación, partiendo de los resultados en un examen básicamente lingüístico, trazando una predicción hacia otras áreas. El examen pragmático, intenta facilitar esta predicción, incluyendo en la prueba ciertos aspectos de la situación comunicativa normal (amplio contexto lingüístico, presiones de tiempo). Sin embargo, ni este tipo de examen se acerca a una situación comunicativa verdadera. Sin estudios cuidadosos de la validez concurrente y/o predictiva de los exámenes de dominio, para comprobar que realmente predicen la actuación en situaciones comunicativas, se tiene que dudar de si estas pruebas pueden dar esta información.

Vollmer en su artículo acerca de dominio general de la lengua

hace más explícito cuál es el interés en exámenes de dominio.

En la evaluación de dominio ... lo que nos interesa principalmente es cómo se aplica el conocimiento lingüístico, cómo se integran los pedazos de conocimientos en diferentes niveles de actuación, cómo se maneja la lengua con más o menos facilidad en términos del efecto comunicativo total de un enunciado.<sup>11</sup>

Y es precisamente acerca de estos elementos que el examen de dominio lingüístico no nos puede dar información concreta.

Se tiene que cuestionar la validez predictiva de los exámenes de dominio lingüístico, preguntando si un examen que sólo mide conocimientos lingüísticos puede informar acerca de la habilidad del examinando para usar los conocimientos para fines comunicativos. Podemos suponer que el individuo que tiene muchos conocimientos de la lengua podría adaptarse a una variedad de situaciones nuevas, como el hablante nativo. Sin embargo, la ausencia de estudios sobre la validez predictiva de los exámenes de dominio ( estudios muy difíciles de llevar a cabo, ya que requieren estudios de seguimiento de individuos en el tiempo, y medidas de criterio posteriores al examen a validar, que sean válidas), justifica, y de hecho obliga la adopción de una actitud crítica y desconfiada ante el poder de instrumentos de dominio lingüístico para dar información acerca de las cosas que señala Vollmer como las importantes.

#### 4.1.6 Examen comunicativo

Ante la inseguridad frente a exámenes de dominio lingüístico, se ha tratado de desarrollar exámenes que den información

acerca del uso de conocimientos lingüísticos para fines comunicativos. Estos exámenes se basan en el análisis teórico de situaciones comunicativas, y en modelos de competencia comunicativa. Tratan de incluir tareas que reflejen de la manera más fiel posible situaciones comunicativas reales.

Morrow, en un artículo sobre la evaluación comunicativa, enumera algunas características del uso de la lengua que no se incluyen en los exámenes de dominio lingüístico. Un instrumento que pretende examinar el uso de la lengua debe contemplar la inclusión de estos elementos de situaciones de uso. Aquí anotamos las características consideradas esenciales por Morrow:

- se basa en la interacción - la producción del examinando depende de lo que se le dice y de las expectativas de su interlocutor.
- se desarrolla de manera no predecible - el examinando tiene que procesar información no predecible con limitaciones normales de tiempo.
- tiene lugar en un contexto - tanto lingüístico como situacional. El examinando debe poder manejar la lengua de manera adecuada a la situación, y al contexto lingüístico.
- todo enunciado tiene un propósito - el examinando debe entender el propósito de los enunciados, y debe poder alcanzar sus propósitos utilizando la lengua.

es auténtico

- los textos para comprensión no deben simplificarse.

la evaluación se basa en resultados concretos

- la comunicación se considera exitosa si logra su propósito extralingüístico. (Morrow nota que este planteamiento es un tanto extremista).<sup>12</sup>

Otro modelo teórico propuesto como base para la elaboración de exámenes de competencia comunicativa es el de Canale y Swain. Ellos dividen la competencia comunicativa en tres partes:

competencia gramatical

incluyendo conocimientos de lexis, reglas de morfología, sintáxis, semántica y fonología.

competencia sociolingüística

incluyendo 1) reglas socioculturales de uso, que permiten producir enunciados de manera apropiado, y comprender la intención de enunciados, y 2) reglas discursivas, cubriendo básicamente cohesión y coherencia.

competencia estratégica

incluyendo estrategias verbales y no verbales para compensar por fallas en la comunicación. Permita al individuo compensar por lagunas en su competencia gramatical, y mantener en general la comunicación.<sup>13</sup>

El modelo de Canale y Swain de hecho ha servido como base por lo menos en un estudio de competencia comunicativa, que incluyó la elaboración de "pruebas" en estas tres áreas.<sup>14</sup>

---

\* Se ha notado que parecería más lógico incluir los conocimientos de cohesión bajo el área de la competencia gramatical. Aquí se respeta la organización original de Canale y Swain.

El modelo más completo y detallado que intenta desglosar los elementos necesarios para poder comunicarse lo constituye el modelo de Munby, específicamente su lista de micro-habilidades.<sup>15</sup> Estas micro-habilidades también han formado la base para la elaboración de exámenes.<sup>16</sup>

El propósito de utilizar un modelo de interacción o competencia comunicativa como base en la elaboración de un examen, es obtener un instrumento que nos dé información acerca de la habilidad del examinando de usar la lengua en situaciones comunicativas. Sin embargo, aquí, al igual que con exámenes de dominio lingüístico, puede haber problemas para generalizar los resultados. Se analizan estos problemas en la siguiente sección.

#### 4.1.7 Examen de actuación versus examen de dominio comunicativo

Un examen que se basa exclusivamente en el establecimiento en el examen de una situación comunicativa, apegándose a las características que enumeró Morrow, sería una evaluación de actuación comunicativa en la situación creada en el examen. Como señala Morrow, esto presenta problemas en la generalizabilidad de los resultados del examen.<sup>17</sup>

El problema de la generalización de los resultados se refiere a la dificultad de llegar a conclusiones acerca de la actuación del candidato en otras situaciones comunicativas no medidas directamente en el examen. En un examen de uso comunica-

tivo oral del inglés para usos sociales, si se evalúa la habilidad del candidato para usar inglés para pedir informes acerca de alojamiento, por ejemplo, ¿cómo podemos saber si el candidato puede pedir informes acerca de trámites gubernamentales (visas, etc.) o establecer y mantener contacto social con hablantes nativos? Generalizar los resultados de pruebas de actuación comunicativa para informar acerca de situaciones no examinadas implica la existencia de una teoría de competencia comunicativa, o de dominio, que postula habilidades y conocimientos que subyacen la actuación comunicativa. En un modelo teórico de este tipo, las habilidades que permiten al candidato actuar en una situación, la situación examinada, le permitirán actuar también en otras situaciones no examinadas.

Las micro-habilidades de Munby ilustran precisamente esta teoría de habilidades subyacentes a la actuación comunicativa, y permiten, por lo menos teóricamente, la elaboración de pruebas de dominio comunicativo o competencia comunicativa, que pretenden medir sistemáticamente no sólo la actuación del examinando en una situación específica, sino el grado en que se domina una serie de habilidades subyacentes a la actuación comunicativa.

En el examen de requisito, se incluyen tareas de actuación, y también secciones que pretenden medir dominio comunicativo, basadas en las micro-habilidades de Munby. De esta manera se espera obtener resultados válidos, y a la vez generalizables.

#### 4.2 Evaluación por criterios versus evaluación por la norma

Aquí se consideran dos maneras de analizar los resultados de un examen; dos maneras de decidir si el resultado de un individuo en un examen se considera "muy bien", "bien", "suficiente" o no aceptable. En el caso de la evaluación por la norma del grupo, esta decisión se toma en base a la relación entre el examinando y otras personas que presentaron el mismo examen. Si un individuo recibió un puntaje muy arriba de la mayoría de los que hicieron el examen se le considera como muy bien, recibe una calificación "alta". En la evaluación por criterios, en cambio, no se compara el resultado del individuo con los resultados del grupo, sino con un resultado ideal preestablecido, basado en el logro de los objetivos que pretende medir el examen. La calificación del individuo depende de si, y hasta qué grado el domina los aspectos que evalúa el examen, y no del grado en que los dominan el grupo en general.

Mientras esta cuestión se refiere en un inicio a la manera en que se interpretan los resultados del individuo en un examen, influye mucho sobre la elaboración del examen. A la persona que elabora un examen con la intención de relacionar los resultados a la norma del grupo, le interesa hacer sobresalir las diferencias que hay entre los miembros del grupo. Quiere saber quién sale arriba o abajo de los demás miembros. Para obtener esta información, es necesario que el examen discrimine, es decir, que dé resultados que diferencien entre los

miembros del grupo a examinar. Existen una serie de pasos a seguir en la elaboración de un examen para lograr que el examen haga resaltar las diferencias entre los examinandos. Entre ellos, hay la eliminación de reactivos que no discriminan, que no demuestran diferencias entre miembros del grupo. Decisiones acerca de la inclusión o no de un reactivo en el examen, entonces, frecuentemente se basan en la manera en que el grupo lo contesta, y no en el contenido del reactivo, lo que está midiendo. En cambio, cuando el examinador quiere saber algo concreto acerca de lo que saben los individuos a examinar, el proceso de elaboración del examen es distinto. Entonces, las decisiones acerca de qué incluir y qué rechazar se basan en el contenido del reactivo o la tarea, y no en la manera en que discrimina entre los miembros de la población a examinar.

La importancia y el valor de los dos procedimientos de evaluación es un asunto que está todavía a discusión. A continuación se consideran las ventajas y desventajas de los dos procedimientos.

Davies afirma que analizar un examen en términos de la norma es un proceso esencial en el desarrollo de un examen, y duda del proceso de análisis con referencia a criterios. Si es que tiene un lugar en el proceso, éste análisis por criterios debe hacerse solamente después de haber establecido el poder discriminatorio del instrumento, según Davies.<sup>18</sup> B. J.

Carroll reconoce la contribución de los dos procesos, e indica que es importante no abandonar todavía el análisis de reactivos, y de resultados de los exámenes, en términos de la norma. Sin embargo, afirma que el "punto principal de referencia" para el diseño del examen y la interpretación de los resultados deben ser los criterios comunicativos.<sup>19</sup> Morrow va más allá cuando aboga por exámenes que rechazan las bases esenciales de la estadística de la norma. Recomienda exámenes que se califiquen cualitativamente, no cuantitativamente, y rechaza la confiabilidad como la característica esencial para obtener una buena medida.<sup>20</sup>

Aquí es esencial volver a señalar la diferencia en el tipo de información que podemos obtener de los dos tipos de exámenes. El examen basado en la norma nos proporciona información acerca de la relación entre los miembros de una población con referencia a un examen. El examen basado en criterios nos dice algo acerca del individuo en relación a criterios que se establecen desde antes de que se escriba el examen.Cuál de los dos tipos de examen servirá mejor, depende del tipo de información que se necesite.

Para ilustrar esto, tomamos y elaboramos un ejemplo que da Carroll para aclarar la distinción entre el criterio y la prueba.<sup>20</sup> El toma como ejemplo un comandante militar que necesita personas que puedan brincar una zanja de doce pies de ancho, para una misión específica. Para seleccionar personal

puede medir directamente si una persona puede cumplir con el mínimo necesario para participar en la misión, haciéndole brincar una zanja de las dimensiones indicadas. Nosotros imaginamos, para fines de comparación, otro comandante que quiere escoger entre todos sus hombres, un grupo especial para misiones difíciles en general. Sin saber las exigencias de todas las futuras misiones, el comandante supone que necesita gente con "buena condición física". Decide que quiere escoger los quince hombres con la mejor condición física para su grupo especial. Para encontrar sus quince hombres, el comandante tendrá que elaborar una serie de pruebas de condición física que le permita ver variaciones que existen en la condición de todos sus hombres. No podrá incluir en su prueba tareas que todos los hombres pueden hacer sin dificultad, ni tampoco tareas que ninguno de sus hombres pueda hacer, sino tendrá que buscar tareas que los mejores sí pueden hacer, y los que tienen menor condición física encuentran imposibles. Es decir, busca pruebas que discriminen en su población, y es coje los quince hombres que tuvieron más éxito en ellas. Lo que hace el segundo comandante es establecer una prueba que le permita ver diferencias que existen en su población, y utiliza los resultados para formar un grupo de los mejores. Es el proceso del examen basado en la norma. El primer comandante, en cambio, establece una prueba que le permite evaluar la capacidad del individuo en relación con un criterio establecido de antemano, y acepta o rechaza a cada individuo en base a su resultado únicamente. Este es el proceso relacionado

con el examen basado en criterios.

Ahora, veamos dos ejemplos relacionados con la evaluación de segundas lenguas. Un encargado de cursos de lengua en una institución quiere establecer tres grupos de estudiantes, uno que contiene los alumnos con más conocimientos, otro con los que tienen menos conocimientos, y otro grupo en medio. Aquí el encargado buscará un examen que mida de manera general los "conocimientos de la lengua", e interpretará los resultados en términos de la norma. En cambio, tomamos otra situación en que el departamento de personal de una compañía de teléfonos debe señalar si los candidatos a trabajos pueden hacer llamadas internacionales en inglés. Aquí, el examinador buscará una medida que se acerque al mayor grado posible al criterio, y analizará los resultados del individuo en términos del criterio.

Con estos ejemplos, vemos la utilidad de los dos tipos de prueba. Cuando se quiere medir si un individuo puede hacer cosas que son posibles de medir de manera más o menos directa (brincar una zanja, hacer una llamada telefónica), se puede trabajar en base a criterios. En cambio, cuando se quiere ver relaciones dentro de un grupo, especialmente en relación con un concepto un tanto vago, y difícil de medir directamente (condición física, conocimientos de la lengua) un examen diseñado para análisis con relación a la norma es la opción que se elegiría.

En estos ejemplos, se ve una distinción muy clara entre situaciones en que sería adecuado el uso de uno u otro de los métodos de análisis de resultados. Desafortunadamente, no siempre es tan claro cuál de los dos métodos sea el adecuado en la práctica. Por una parte, puede haber confusión en relación a si se quiere información acerca del grupo, o del individuo. También puede haber confusión en cuanto a lo que se está evaluando, y si se presta o no realmente a la medición directa. A veces podemos encontrar que se utilizan resultados que reflejan el análisis de los resultados en términos de la norma, para hacer decisiones acerca de individuos, decisiones que implican la aplicación de criterios. Un ejemplo de esta situación es el uso de resultados en el examen TOEFL para decidir si un candidato tiene el inglés necesario para estudiar un programa académico en inglés. Aquí se está suponiendo que haber recibido una puntuación, por ejemplo de quinientos cincuenta significa que el candidato sí podrá llevar a cabo sus estudios en inglés. Pero el diseño del examen no justifica este juicio (a menos de que se hayan llevado a cabo estudios para establecer la validez predictiva del examen).

Otro factor importante que introduce confusiones cuando se diseña un examen es el peso de la tradición en la evaluación que da valor positivo a una serie de características de un buen examen para discriminar entre miembros de un grupo (buena discriminación, ni muy fácil ni muy difícil, etc.) que son irrelevantes o hasta contraproducentes para la elaboración de

un examen basado en criterios. Esta tradición también proporciona un cuerpo de procedimientos para establecer el poder discriminatorio de los reactivos, y el buen funcionamiento de ellos. Como señala Ingram hay cierta falta de técnicas para establecer el buen funcionamiento de reactivos en exámenes evaluados en términos de criterios.<sup>22</sup> Sin embargo, se están desarrollando procedimientos para analizar el funcionamiento de reactivos en exámenes basados en criterios. Carroll presenta un sistema en que se comparan los resultados de diferentes grupos de examinandos en reactivos particulares.<sup>23</sup>

#### 4.3 Confiabilidad

Un examen es confiable siempre y cuando arroje información estable acerca del examinando. Es decir, si un examen se aplica dos veces a la misma población con un lapso de una semana entre las dos aplicaciones, y se encuentran grandes diferencias en los resultados de los individuos, entre la primera y la segunda aplicación, el examen no puede considerarse como confiable. Tampoco será confiable el examen que, corregido por dos examinadores, da diferentes resultados.

#### 4.4 Validez

La validez de un procedimiento de evaluación (el examen y las condiciones en que se aplica y se califica), se refiere al grado en que éste mide lo que pretende medir, y no otras cosas. En el caso de una medida directa, la cuestión de validez pierde importancia. Retomando el ejemplo de la prueba

para seleccionar personas que pueden brincar una zanja con doce pies de ancho\* podemos ver claramente que la prueba directa sí mide precisamente lo que se quiere medir. Pero como ya se mencionó, casi nunca es posible obtener una medida directa debido a la naturaleza del objeto a medir. Generalmente, es imposible usar una medida directa por varias razones, entre ellos: 1) lo que se quiere medir es tan grande, tan extenso que no sería práctico administrar un examen que cubriera la totalidad del objeto a evaluar, y 2) lo que se quiere medir no es observable directamente, entonces hay dificultad en definir exactamente qué es.

El primer caso se daría si se quisiera saber si el examinando puede dar el equivalente de las palabras del inglés, en su lengua materna. Aquí "las palabras del inglés" representa un número muy grande de elementos a evaluar, y sería muy poco práctico examinarlos todos. El examen que se elaboraría, entonces, no podría pretender medir el objeto de manera directa; se tendría que seleccionar del universo de objetos, un muestreo a medir. El segundo problema mencionado es el de los exámenes de dominio. Como no se puede observar de manera directa el "dominio" del candidato, y no hay una definición aceptada y completa de lo que significa dominar una lengua, no es posible medirlo de manera directa. Se tiene que hacer un instrumento basado en una definición del concepto "dominio"

---

\* Véase la sección 4.2

y esta definición operacional puede ser un reflejo más o menos fiel de la realidad. Generalmente, estos dos problemas van juntos. En el primer ejemplo, hemos definido el examen como de traducción de palabras. Si realmente se quisiera saber si el examinando puede traducir palabras, bien. Pero si se quiere medir si el examinando "sabe" las palabras del inglés, el problema se complica, por la falta de una definición clara de lo que se quiere decir por "saber" las palabras. Así se juntan los dos problemas; la cantidad de elementos a medir, y la falta de precisión acerca de lo que se quiere medir. A la hora de tratar de establecer la validez de un examen, se consideran estas dos maneras en que la prueba difiere de una medida directa.

La determinación de la validez del contenido se relaciona con el problema del muestreo, el hecho de que no sería posible medir directamente todas las cosas que se quiere saber. El documento "Standards for Educational and Psychological Measurement" establece la manera de demostrar la validez de contenido:

Para demostrar la validez de contenido de un grupo de resultados de una prueba, se tiene que demostrar que el comportamiento demostrado en el examen constituye una muestra representativa del comportamiento a demostrar en el dominio de actuación deseado. La definición del dominio de actuación, los objetivos del usuario, y el método de hacer la muestra, son elementos esenciales para demostrar la validez de contenido.<sup>24</sup>

Entonces, en el grado en que el contenido de un examen no es la totalidad de lo que se quiere medir, será necesario justi-

ficar la muestra que se mide, definiendo el universo a evaluar e indicando la manera utilizada para llegar a la muestra evaluada. En este sentido, la validez de contenido es importante para cualquier medida no absolutamente directa y completa. Es bastante sencillo establecer la validez de contenido en exámenes de aprovechamiento, en que el universo a medir está comodamente definido en el programa del curso. En exámenes de otro tipo, de dominio por ejemplo, el problema es mucho más complicado, pero no por esto de menor importancia.

La dificultad en definir el objeto a medir, por ser una cosa no directamente observable, lleva a la necesidad de establecer la validez de construcción del examen. La comprensión de lectura sirve de ejemplo. Mientras tal vez todos entienden más o menos lo que se quiere decir con "comprensión de lectura", no existen respuestas finales y universalmente aceptadas a preguntas como ¿qué es la lectura? ¿qué significa leer? ¿qué significa comprender lo que se lee? y ¿cómo se puede saber que alguien haya comprendido lo que ha leído? Mientras que estas preguntas pueden parecer bastante teóricas, y el hecho de elaborar exámenes una cosa práctica, no existe una división aquí entre la teoría y la práctica. De hecho, cada examen, sea de comprensión de lectura, de dominio o de comprensión auditiva, es la operación de una teoría, o sea implícita o explícita, acerca de la cosa que se quiere medir. Un examen de dominio constituido por cien reactivos de puntos aislados yace en la teoría que dominar el idioma es dominar

una serie de puntos aislados. Entonces, evaluar la validez de la construcción de un examen implicaría una definición explícita de la teoría subyacente al examen, y la demostración que es una teoría válida.

Cuando la teoría detrás del examen se refiere esencialmente al sistema lingüístico, la validez de construcción se establece en base a la teoría lingüística. Los exámenes de dominio de reactivos de puntos aislados, que evalúan reconocimiento de estructuras, se basan en la teoría de la lingüística estructural. Sin embargo, todo examen de hecho refleja no sólo una teoría acerca de la naturaleza del objeto a evaluar, sino también una teoría acerca de qué significa saberlo. Aquí entra la teoría psicolingüística, y la necesidad de investigar no sólo el contenido del examen, sino la realidad psicolingüística de las teorías subyacentes al examen. Para hacer esto se puede recurrir a una serie de pruebas experimentales y el análisis estadístico de los resultados para comprobar la realidad psicolingüística del aspecto que se quiere medir. Se revisan y se discuten algunos estudios de este tipo en la siguiente sección (4.5) de este trabajo.

Los dos acercamientos a la validez mencionados hasta ahora (del contenido y de la construcción) enfocan la atención sobre el examen a validar, asegurando que su contenido refleja lo que se quiere medir, y que se conforma a una teoría válida de qué significa saber lo que se está midiendo.

Otros dos acercamientos a la validación enfocan no el instrumento a validar, sino un criterio independiente. Estos dos enfoques evalúan el grado en que el instrumento a validar da los mismos resultados que otras medidas.

Se establece la validez concurrente si se obtienen correlaciones altas con otra medida aplicada en el mismo momento que el examen a validar. Aquí no interesa, entonces, el contenido del instrumento a validar; lo importante es la relación entre los resultados obtenidos en ésta medida, y otra medida válida de lo que se pretende medir. Así que sería posible que resultara tener alta validez concurrente un examen integrativo con un examen de puntos aislados. La dificultad en establecer la validez concurrente es que se necesita encontrar una medida válida con la cual comparar los resultados del instrumento a validar. Si no existe tal medida, no será posible establecer la validez concurrente de un instrumento nuevo.

La validez predictiva se refiere al grado en que los resultados de un examen reflejan lo que será la actuación futura del examinando, en un momento posterior a la evaluación. Este tipo de validez es esencial ya que el uso de muchos exámenes presupone una validez predictiva que a veces nunca se ha comprobado. Si un examen se utiliza para la selección de personal, por ejemplo, se está suponiendo que este examen predice el grado en que el candidato podrá desempeñar las funciones del trabajo futuro. Si el examen no hace esto, se vuelve

requisito burocrático, o injusta e innecesaria barrera.

La dificultad de establecer la validez predictiva de una medida se relaciona con el paso de tiempo entre la aplicación del examen y la recepción de datos que harán posible una comparación. Por una parte, este lapso de tiempo implica una larga espera antes de tener información acerca de la validez predictiva de un instrumento nuevo. Por otra parte, frecuentemente se pierden los examinandos de vista, y no se obtienen los datos suficientes para hacer la comparación. Sin embargo, a pesar de las dificultades, es necesario intentar establecer la validez predictiva, sobre todo cuando los resultados del examen de hecho se utilizan para hacer predicciones acerca de la actuación futura del examinando.

Es de notar que el proceso de establecer la validez de un examen es difícil e interminable. No es posible hablar en términos absolutos, diciendo que un examen "es válido". Un instrumento aplicado a una población determinada, de manera determinada, puede tener más o menos validez para uno u otro propósito. Y los estudios de validez que se lleven a cabo con cualquier examen (de contenido, de construcción, concurrente o predictiva) depende del propósito del examen y el uso que se le dé.

#### 4.5 Qué evaluar? Consideraciones sobre la validez de construcción

Una de las preguntas centrales presentes en la investigación en la evaluación se relaciona con las divisiones que se pueden hacer en el objeto a evaluar. Esta pregunta es central no sólo a la investigación de la hipótesis de competencia única,<sup>25</sup> sino a la elaboración de todo examen de dominio y de lenguas para propósitos específicos.

En el caso del diseño del examen de requisito, la validez de construcción ha tomado el interés principal. Se quería lograr un examen que midiera el dominio comunicativo de los candidatos. Por lo tanto desde el diseño, era necesario explorar el concepto de dominio comunicativo, para llegar a una definición operacional, que diera validez de construcción al examen. En la siguiente sección, se explora el concepto de dominio lingüístico, revisando algunos estudios de la validez de exámenes que intentan medirlo (sección 4.5.1). Luego se revisan estudios de la construcción "competencia comunicativa" (sección 4.5.2).

##### 4.5.1 Estudios de la construcción "dominio lingüístico"

Como se ha dicho, un examen de dominio general de la lengua pretende medir algún factor "L", conocimientos de la lengua. El criterio para la selección del contenido de este tipo de examen es lingüístico exclusivamente; no se basa en ningún análisis del examinando y sus usos futuros del idioma.

Establecer exactamente qué mide (o qué quiere medir) un examen de dominio lingüístico es un problema difícil. Afirma Vollmer que "el dominio lingüístico es lo que miden los exámenes de dominio lingüístico",<sup>26</sup> para señalar el grado de falta de conocimientos en este renglón.

Frecuentemente, los exámenes de dominio han incluido una serie de subpruebas para medir diferentes aspectos de dominio, como son: vocabulario, gramática, comprensión auditiva, comprensión de lectura, producción oral, producción escrita, etc. Implícitamente, estos exámenes se basan en una definición de dominio de lengua que incluye una serie de dominios en diferentes áreas. A la hora de investigar más a fondo la construcción de dominio lingüístico, se ha cuestionado la validez teórica de dividir el dominio en los elementos tradicionalmente evaluados. Investigaciones de los últimos diez años han intentado definir claramente la construcción "dominio" de una lengua, y a la vez, de validar exámenes que pretenden medirlo.

El trabajo de Oller, en particular ha tenido el propósito de definir y demostrar la realidad psicolingüística de la construcción teórica que define el dominio de la lengua como una competencia única que subyace toda actuación lingüística. Los exámenes integrativo-pragmáticos, según Oller, permiten medir esta competencia única. En este tipo de examen, el examinando se enfrenta a un texto que contiene las característi-

cas contextuales del discurso extendido, y debe poderlos manejar, bajo restricciones de tiempo. Y debe comprender el texto a nivel semántico, y no sólo sintáctico. En las palabras de Oller, "...las secuencias lingüísticas deben relacionarse con contextos extra-lingüísticos de manera significativa."<sup>27</sup>

Oller y otros han hecho estudios de la validez de construcción de los exámenes pragmáticos, sobre todo el dictado y el cloze, y han concluido que estos exámenes sí logran evaluar de manera eficiente el dominio que el examinando tiene de la lengua. Han encontrado correlaciones altas entre estos exámenes y otros que tratan de evaluar habilidades aisladas como la comprensión auditiva, y de lectura, etc. Todo esto demuestra para Oller que todos estos exámenes miden una sola cosa, la existencia en el examinando de una "gramática de expectativas",\* su competencia única.<sup>28</sup>

La evidencia que apoya el punto de vista de Oller es mucha; sin embargo, la veracidad de la teoría de competencia única no se ha establecido. Vollmer reporta que el mismo Oller, en un simposio sobre la evaluación en 1980, "... admitió públicamente que tal vez se había pasado en su postulación de una só la habilidad indivisible, subyacente a todas las diferentes

---

\*Esta competencia incluye la capacidad de predecir en base a todo un contexto léxico, semántico y sintáctico. No se refiere únicamente al sistema gramatical.

actividades lingüísticas en que un aprendizando pueda participar."<sup>29</sup> De hecho, la cuestión del grado de divisibilidad de habilidades comprobable, sigue en pie.

#### 4.5.2 Estudios de la construcción "competencia comunicativa"

En los últimos años, y en gran parte en respuesta al trabajo de Oller, se ha despertado más interés en el establecimiento de la validez de construcción. Se está haciendo una serie de estudios <sup>30</sup> en que la preocupación ha sido, supuestamente, establecer la validez (o no-validez) de dividir la competencia comunicativa en diferentes habilidades (competencia comunicativa para la lectura en contraste con la competencia para la producción oral, por ejemplo), o en diferentes áreas de la competencia comunicativa (competencia lingüística, competencia socio lingüística y estratégica, según el modelo de Canale y Swaine<sup>31</sup> por ejemplo). Se ha tratado de establecer la realidad psicolingüística de estas divisiones de la competencia comunicativa, y entonces, la validez de construcción de exámenes que reflejan estas divisiones. Aunque algunos de estos estudios han seguido un modelo de competencia comunicativa,<sup>32</sup> otros se basan en exámenes de dominio que no se diseñaron en base a un modelo comunicativo<sup>34</sup> y por lo tanto no investigan realmente la construcción "competencia comunicativa". Los estudios de Palmer y Clifford son parte de un proyecto de investigación de validez de construcción que nació en dos coloquios sobre evaluación que tuvieron lugar en las convenciones de TESOL en 1979 y 1980. El volumen The Construct Validation of

Tests of Communicative Competence, editado por Palmer, Groot y Trosper, y publicado por TESOL informa sobre el proyecto. En los estudios que se refieren a la validez de construcción, (algunos tratan confiabilidad, validez de contenido, y validez concurrente), la preocupación ha sido con la pureza de la medida. Es decir, si se quiere medir, por ejemplo, la competencia comunicativa para la lectura, se debe medir esto y nada más. Se debe establecer que existe una habilidad especial para comprensión de lectura, independiente de otros factores, como podrían ser conocimientos gramaticales, vocabulario, etc. Si resulta que todas las diferencias que demuestran los examinandos en su competencia para lectura se deben a diferencias en sus conocimientos gramaticales, y su vocabulario (establecido por medio de pruebas independientes de estos factores), entonces no se puede afirmar que hay una habilidad específica que es competencia en comprensión de lectura, y en un sentido estricto, no tendrá validez de construcción el examen de lectura.

Cabe mencionar que la aplicación de esta interpretación de validez de construcción no es universal en el área de evaluación en segundas lenguas. Han sido Palmer y otros participantes en los coloquios de TESOL quienes se han preocupado por la validez de construcción en este sentido, de exámenes de comunicación oral. Para establecer la pureza del examen (y al mismo tiempo la independencia en teoría de la habilidad que el examen mide), se tendría que asegurar que factores externos

a la habilidad o construcción meta no influyen en los resultados. Para comprobar esto, Stevenson ha sugerido la aplicación de una técnica de validación propia de pruebas psicológicas: una metodología convergente-discriminante.<sup>34</sup> Esta metodología pretende establecer si dos medidas diferentes que pretenden medir la misma cosa, pero con métodos diferentes, sí dan resultados parecidos, y si dos medidas que pretenden medir cosas distintas, utilizando la misma metodología, dan resultados diferentes.

Para hacer un experimento de este tipo, se hacen dos pruebas de la misma habilidad, o del mismo aspecto, utilizando dos formatos distintos (opción múltiple y llenar espacios, por ejemplo), y dos pruebas de otra habilidad, utilizando los dos formatos usados para la otra habilidad. Se establece así, una matriz:

		formato	
		X	Y
habilidad	A	(1) habilidad A, formato X	(2) habilidad A formato Y
	B	(3) habilidad B formato X	(4) habilidad B formato Y

CUADRO 4-1

MATRIZ PARA UN ESTUDIO CONVERGENTE-DISCRIMINANTE

Para establecer la independencia de las habilidades "A" y "B" se deben obtener correlaciones entre pruebas 1 y 2, y entre 3 y 4 que sean superiores a las correlaciones entre las pruebas 1 y 3, y 2 y 4. Esto demostraría que lo que más influye

en los resultados es la habilidad o el aspecto a medir, y no el método utilizado para medirlo.

Estos procedimientos permiten distinguir entre la variación en los resultados que es producto de diferencias en la habilidad a medir, y la variación que depende del llamado "factor método". El término "factor método" se refiere al hecho de que en cualquier medida indirecta, la manera en que se hacen las preguntas (el método), y no sólo el contenido de ellos (la habilidad o aspecto a medir) influye en los resultados. Esta técnica de validación trata de valorar y eliminar la influencia del método, para ver claramente si el examen mide la habilidad o construcción deseada. Si los resultados demuestran que el factor método es más fuerte que el factor habilidad o elemento a medir, puede ser por una de dos razones. Por una parte puede ser porque los exámenes no son buenos, y no permiten ver las diferencias que realmente existen en los examinados entre una y otra habilidad. Por otra parte, puede ser porque las dos habilidades no son realmente independientes: que habilidad "A" y habilidad "B" dependen de, o son manifestaciones de una sola habilidad "C" que subyace a ellas.

Esta metodología para validación se ha aplicado en el análisis de los resultados de algunos exámenes. En un estudio, se analizaron las relaciones entre las subpruebas de Lectura, Producción Escrita, Producción Oral, y Comprensión Auditiva en dos medidas: un examen estandarizado (del "Modern Language

Asociation" en alemán), y la evaluación del maestro del examinando.<sup>35</sup> El análisis de los datos en este estudio no permitió ver las cuatro subpruebas como medidas de habilidades distintas, ya que en los resultados de las subpruebas, influyó más el método de evaluación que la habilidad evaluada.

Esta conclusión no es sorprendente, sobre todo tomando en cuenta la masa considerable de datos que ha presentado Oller, argumentando en favor de la hipótesis de la competencia única en la lengua. Es fácil aceptar que la división tradicional en cuatro habilidades no muestra de manera fiel la realidad psicolingüística del conocimiento de la lengua, reflejada en estas pruebas.

Es importante notar que en las discusiones de los estudios de validez de construcción en que se utiliza la metodología convergente-discriminante, la atención se enfoca mucho más en los procedimientos de validación que en la construcción a medir. Mientras que Stevenson nota la falta general de análisis y especificación de las situaciones que el examen pretende predecir o reflejar,<sup>36</sup> en la colección en que se encuentra su artículo, el énfasis cae casi totalmente en el proceso de validación y no en la construcción a medir.<sup>37</sup> Poca atención se da a la necesidad de definir con la mayor claridad posible, dado el estado de la teoría y análisis lingüísticos, lo que se considera son los componentes esenciales a evaluar. La excepción a esta falta de trabajo analítico y teórico en este volumen es una

contribución breve de Canale y Swain en que se postula un modelo de competencia comunicativa que incluye los niveles de competencia lingüística, sociolingüística y estratégica.\*

Mencionan el trabajo de Munby como una probable base para detallar su modelo.<sup>38</sup> Sin embargo, ninguno de los estudios experimentales intenta un análisis o descripción detallada de lo que los exámenes investigados pretenden medir. En el estudio de Clifford arriba mencionado, el autor se contenta con aceptar las etiquetas que la "Modern Language Association" dió a sus pruebas. La preocupación con los procedimientos para la validación resta atención al examen mismo y a lo que pretende medir.

Este acercamiento a la validez de un examen se caracteriza por la aceptación de la validez técnica, estadística del examen como el aspecto más importante; el énfasis primario en las técnicas de validación estadística, y no en la definición cuidadosa de la construcción, y la elaboración de exámenes en base a ella. Esta actitud hacia la validez no es universal, y el trabajo de Morrow, Carroll, y otros refleja otro acercamiento a la validez, y a la elaboración de exámenes en general.

Algunas diferencias entre estas dos "escuelas" se verán claramente por medio de unas citas de las dos. Primero encontramos:

---

\* Véase la sección 4.1.6.

Adoptar una prueba por el simple hecho de que se ve aceptable es una mala idea; muchas pruebas que se ven bien han fallado como indicadores ... este tipo de evidencia (reforzado por toda la historia de la frenología, grafología y pruebas de brujería) sirve de precaución en contra de la adopción de una prueba simplemente porque se ve bien. Si hay que escoger entre una prueba con 'validez visual' y sin validez técnica verificada, y otra con validez técnica, pero sin atractivo para el individuo no experimentado, se debe escoger esta última. 39

[La validez visual] no es [un atributo] que un especialista de la medición ofrecería a otro como evidencia de que su medida refleja, de hecho, una construcción extremadamente compleja.<sup>40</sup>

La validez visual puede considerarse como la apariencia de la validez, de la perspectiva de un observador no experimentado en la medición.<sup>41</sup>

Sin embargo, existen otras opiniones al respecto:

A pesar de lo que pueden pensar Palmer y Bachman, la validez visual es de gran importancia cuando una prueba se administra por personas no orientadas hacia la estadística.<sup>42</sup>

Esta otra actitud ante la validez técnica, se resume en esta cita:

Morrow enfatiza correctamente que, aunque las medidas indirectas de habilidades lingüísticas dicen tener niveles extremadamente altos de validez y confiabilidad, establecidos por procedimientos estadísticos, debemos sospechar de esta validez.<sup>43</sup>

Se nota una actitud de desconfianza ante las técnicas de validación consideradas por otros como esenciales, necesarias, y como las únicas que pueden dar información verídica acerca de las medidas. Y por otra parte, se nota un respeto y reconocimiento de la importancia de la validez visual, aspecto que recibe desprecios y bromas de otros, cuando se comparan

exámenes que la tienen con "pruebas de brujería".

Esto refleja dos acercamientos distintos a la validación, y sobre todo a la validación de construcción Morrow y Carroll, en lugar de basar estudios de validez sobre el análisis de resultados de exámenes ya existentes (como lo ha hecho Clifford) tratan desde un inicio de elaborar exámenes válidos, desde el punto de vista de la construcción que pretenden medir. Esto se logra mediante el trabajo teórico de hacer explícito qué es lo que se quiere medir realmente; de definir en qué consiste la habilidad, elemento, la construcción a medir. En el momento en que se haya hecho explícita la construcción, en términos de un modelo teórico, se le puede hacer operacional en un examen que será congruente con este modelo.

En relación con este acercamiento al problema de la validez del examen, encontramos otra definición del proceso de validación de construcción. Aquí se hace un análisis cuidadoso del examen, sus sub-pruebas y reactivos, para averiguar si realmente reflejan la construcción que se pretende medir. Esta definición de validez de construcción está clara y explícita en el trabajo de Carroll<sup>44</sup> e implícita en otros tratamientos del tema.<sup>45</sup> Este proceso, a nuestro parecer, haría falta en cualquier estudio de validez de construcción, pero está ausente en algunos estudios que utilizan la metodología rigurosa convergente-discriminante.

Aunque esta segunda manera de hacer la validación de la construcción no pretende comprobar empíricamente si la construcción en que se basa el examen refleja una realidad psicolingüística, sí pretende asegurar que el examen refleje de manera fiel una teoría bien articulada, una teoría que refleje el "estado del arte" en teoría y análisis lingüísticos, y que satisfaga al autor del examen. El estudio de validez convergente-discriminante puede terminar dando información acerca de si lo que mide el examen estudiado refleja una realidad psicolingüística, pero no dará información acerca de ninguna construcción teórica si el examen no se hizo con base en alguna teoría.

Desde el punto de vista práctico, parecería mucho más importante, y de hecho sería el primer paso necesario, elaborar exámenes que reflejen de la manera más fiel posible lo que el examinador considera ser lo esencial de lo que quiere medir. Tal vez es aquí de donde vienen discrepancias acerca de la importancia de la validez visual. Algunos definen la validez visual como la opinión de una persona sin experiencia en la evaluación. Pero para otros, la validez visual de una subprueba puede ser la opinión de una persona muy experimentada en la evaluación que dice "a mí me parece que esta tarea refleja lo que yo he decidido que es importante medir, después de larga consideración de la construcción a evaluar."

## Notas de referencia bibliográfica

### Capítulo IV

<sup>1</sup>Alan Davies, "Oral English Testing in West Africa," Language Testing Symposium: A Psycholinguistic Approach, ed. Alan Davies. (London: Oxford University Press, 1968), p. 79.

<sup>2</sup>Rebecca Valette, Modern Language Testing, 2nd ed. (New York: Harcourt, Brace, Jovanovich, 1977).

<sup>3</sup>John W. Oller, "Discrete-Point Tests versus Tests of Integrative Skills," en Focus on the Learner: Pragmatic Perspectives for the Language Learner, eds. John Oller y Jack C. Richards, (Rowley: Newbury House Publishers, 1973), p. 190.

<sup>4</sup>John W. Oller, Language Tests at School: A Pragmatic Approach (London: Longman, 1979), p. 37.

<sup>5</sup>Ibid., pp. 38-39.

<sup>6</sup>Renate Schultz, "Discrete Point versus Simulated Communication Testing in Foreign Languages," The Modern Language Journal LXI, 3 (1977): 96.

<sup>7</sup>Alan Davies, "The Construction of Language Tests," en The Edinburgh Course in Applied Linguistics Volume 4 Testing and Experimental Methods, eds. J. P. B. Allen y Alan Davies (Oxford: Oxford University Press, 1977), p. 50.

<sup>8</sup>A. E. G. Pilliner, "Subjective and Objective Testing," en Language Testing Symposium, ed. Alan Davies, p. 20.

<sup>9</sup>Jonathon Anderson, Psycholinguistic Experiments in Foreign Language Testing (St. Lucia, Queensland: University of Queensland Press, 1976), pp. 142-143.

<sup>10</sup>Alan Davies, "The Construction of Language Tests," p. 46.

<sup>11</sup>Helmut J. Vollmer, "Why are we interested in 'General Language Proficiency'?" en Issues in Language Testing, eds. J. Charles Alderson y Arthur Hughes (London: The British Council, 1981), p. 160.

<sup>12</sup>Keith Morrow, "Communicative language testing: revolution or evolution?," en The Communicative Approach to Language Teaching, eds. C. J. Brumfit y K. Johnson. (Oxford: Oxford University Press, 1979), pp. 149-150.

<sup>13</sup>Michael Canale y Merrill Swain, "A theoretical framework for communicative competence," en The Construct Validation of Tests of Communicative Competence, eds. Adrian S. Palmer; Peter J. M. Groot; y George A. Trostler (Washington, D. C.: TESOL, 1981), pp. 31-36.

<sup>14</sup>L.F. Bachman y A.S. Palmer, "The Construct Validation of Tests of Communicative Competence," Trabajo presentado en la 16a Convención de TESOL, 1-6 Mayo, Honolulu, Hawaii.

<sup>15</sup>John Munby, Communicative Syllabus Design: A sociolinguistic model for defining the content of purpose specific language programmes. (Cambridge: Cambridge University Press, 1978), 123-131.

<sup>16</sup>Brandon J. Carroll, Testing Communicative Performance: An Interim Study (Oxford: Pergamon Press, 1980).

<sup>17</sup>Keith Morrow, "Communicative Language Testing," p. 151.

<sup>18</sup>Alan Davies, "Introduction," p. 8.

<sup>19</sup>Brendon J. Carroll, Testing Communicative Performance, p. 10.

<sup>20</sup>Keith Morrow, "Communicative Language Testing"

<sup>21</sup>Brendon J. Carroll, Testing Communicative Performance, p. 101.

<sup>22</sup>Elizabeth Ingram, "Basic Concepts in Testing," en Testing and Experimental Methods, eds. Allen y Davies, p. 32.

<sup>23</sup>Brendon J. Carroll, Testing Communicative Performance, pp. 65-66.

<sup>24</sup>American Psychological Association, "Standards for educational and psychological tests," en Construct Validation, eds., Palmer, et al., p. 42.

<sup>25</sup>Véase John W. Oller, Language Tests at School; John W. Oller y Francis Butler Hinofotis, "Two mutually exclusive hypotheses about second language ability: indivisible or partially divisible competence," en Research in Language Testing, eds. John W. Oller y Kyle Perkins (Rowley: Newbury House Publishers, 1980); John W. Oller y Kyle Perkins, eds. Language in Education testing the tests (Rowley. Newbury, 1978).

- 26 Helmut J. Vollmer, "Why are we interested in 'General Language Proficiency'?", p. 152.
- 27 John W. Oller, Language Tests at School, pp. 38-39.
- 28 Ibid.
- 29 Helmut J. Vollmer, "Response: Issue or non-issue: General Language Proficiency Revisited," en Issues in Language Testing, eds. Alderson y Hughes, p. 196.
- 30 Adrian S. Palmer, Peter J. M. Groot y George A. Trostler, eds., The Construct Validation of Tests of Communicative Competence (Washington, D. C.: TESOL, 1981).
- 31 Michael Canale y Merrill Swain, "Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing," Applied Linguistics 11.
- 32 Lyle F. Bachman y Adrian S. Palmer, "The Construct Validation of Tests of Communicative Competence," trabajo presentado en la 16a convención anual de TESOL, Honolulu, Hawaii, 1-6 mayo, 1982.
- 33 Ray T. Clifford, "Convergent and discriminant validation of integrated and unitary language skills: the need for a research model," en Construct Validation, eds., Palmer, et al.
- 34 Douglas K. Stevenson, "Beyond Faith and Face Validity: the multitrait-multimethod matrix and the convergent and discriminant validity of oral proficiency tests," en Construct Validation, eds., Palmer, et al., p.
- 35 Ray T. Clifford, "Convergent and discriminant validation"
- 36 Douglas K. Stevenson, "Beyond faith and face validity," p. 43.
- 37 Adrian S. Palmer, et al., eds., Construct Validation.
- 38 Michael Canale y Merrill Swain, "A theoretical framework for communicative competence," p. 34.
- 39 L. J. Cronbach, Essentials of psychological testing 3rd ed. (New York: Harper & Row), p. 183, citado en Construct Validation, eds. Palmer, et al., p. 42.

<sup>40</sup>Stevenson, "Beyond faith and face validity", p. 40.

<sup>41</sup>Ibid., p. 41.

<sup>42</sup>Carol M. Clapham, "Reaction to the Carroll Paper (1)," en Issues in Language Testing, eds., Alderson y Hughes, p. lll.

<sup>43</sup>Cyril J. Weir, "Reaction to the Morrow Paper (1)," en Issues in Language Testing, eds., Alderson y Hughes.

<sup>44</sup>Carroll, Testing Communicative Performance, p. 67.

<sup>45</sup>Morrow, "Communicative language testing: revolution or evolution."

## CAPITULO V

### EL EXAMEN DE REQUISITO PARA UN SISTEMA DE PREPARACION PARA BECARIOS

#### 5.0 Introducción

Tomando en cuenta los temas de evaluación resumidos antes, se buscó un diseño para el examen de requisito para becarios. Aparte de las consideraciones teóricas de la evaluación, se tuvieron en mente algunas condiciones que el examen tendría que cumplir para ser un instrumento útil en la práctica. Se necesitaría:

- 1) Un diseño flexible, aplicable en diferentes situaciones con diferentes recursos disponibles
- 2) Un instrumento económico que maximizara la información obtenida, con el mínimo de esfuerzos necesarios para aplicar el examen y evaluar los resultados. Tomamos como supuesto, sin embargo, que sería imprescindible la participación de un examinador preparado.
- 3) Un instrumento que arrojará información detallada, pero comprensible y útil acerca de cada candidato, para el uso de organismos que otorgan becas, y deben tomar decisiones acerca de los candidatos.
- 4) Un instrumento que tuviera, idealmente, un alto grado de validez predictiva, de acuerdo al propósito para el cual había sido diseñado.

Como enfoque general al examen de requisito, hemos tomado el principio de utilizar una variedad de tipos de medidas para obtener una idea más completa acerca del candidato en su conocimiento y uso de la lengua. Este enfoque rechaza la idea de que es posible encontrar una prueba (o un tipo de prueba) que dé la información deseada.

Se tomó este punto de vista por dos razones. Por una parte, no hay datos suficientes para asegurarnos que lo que nos interesa se relaciona con un sólo factor de dominio lingüístico, o uso comunicativo. Y por otra parte, como señalan Volmer<sup>1</sup> y Alderson,<sup>2</sup> dada la falta de seguridad en cuanto a las medidas de evaluación, y la importancia de las decisiones que se hacen en base a los resultados, para ser justos con los examinandos, deben diversificarse y ampliarse las medidas.

Por estas razones, el diseño refleja no una sola visión o teoría coherente, sino varios acercamientos al problema de la definición y medición del conocimiento y uso de la lengua.

En el cuadro 5-1 se presenta un resumen del diseño del examen, indicando las construcciones a medir, la forma que toman las partes del examen, con indicaciones acerca de su calificación (objetiva o subjetiva), el análisis de los resultados (en relación con la norma o criterio), y el tipo de validez que se pretende alcanzar en cada parte.

N I V E L	L I N G Ü I S T I C O		C O M U N I C A T I V O	
	Dominio lingüístico		Micro-habilidades	Actuación comunicativa
CONSTRUCCION A MEDIR	Factor "L" - conocimientos lingüísticos		Una serie de micro-habilidades subyacentes a la actuación comunicativa	Prueba directa de actuación comunicativa
FORMA QUE TOMA LA(S) PRUEBA(S)	Puntos aislados	Integra- tivo	Tareas que activan micro-habilidades particulares de manera relativamente aislada	Tareas comunicativas globales, que implican el uso en conjunto de conocimientos lingüísticos y micro-habilidades
CALIFICACION	Objetiva		Objetiva/subjetiva	Subjetiva
ANALISIS DE LOS RESULTA- DOS	Norma/criterio		Criterio	
VALIDEZ	P r e d i c t i v a			
	Concurrente construcción		Construcción contenido	Visual/contenido

CUADRO 5-1

DISEÑO GLOBAL DEL EXAMEN DE REQUISITO PARA BECARIOS

### 5.1 Evaluación del nivel lingüístico

Se incluye una parte del examen en que se evalúa el nivel de conocimientos que tiene el examinando acerca de la lengua, sin investigar su aplicación en una situación comunicativa. Mientras que es claro que este tipo de conocimientos de la lengua constituye sólo un elemento entre varios que operan en el uso comunicativo de la lengua, es un elemento esencial sin el cual sería imposible echar a andar estrategias comunicativas, y conocimientos sociológicos del contexto, a manera marginalmente satisfactoria, sobre todo considerando la complejidad de las tareas que tendrá que cumplir el becario en su curso académico. La "competencia lingüística" es uno de los tres niveles que postulan Canale y Swain en su modelo de competencia comunicativa, y varias de las micro-habilidades de Munby se refieren a conocimientos del sistema gramatical.

El área de dominio lingüístico se divide en dos en el diseño del examen: una parte que mide puntos aislados de la gramática, y otra parte integrativa y pragmática.

La sección de puntos aislados debe enfocar el sistema gramatical de la lengua, tomando éste, en el grado en que sea posible, en su aspecto nocional. Es decir, el énfasis sería no sólo en la forma o estructura gramatical, sino también en el uso del sistema para expresar nociones. Se incluyen el significado de los diferentes tiempos verbales, la expresión de nociones de cantidad y comparabilidad, etc.

La importancia de la evaluación del aspecto gramatical yace en el hecho de que este sistema es poderosamente generativo, y permite al individuo mostrar flexibilidad y capacidad para adaptarse ante situaciones, siempre y cuando posea los conocimientos sociolingüísticos necesarios, y estrategias comunicativas.

Una parte del examen de dominio lingüístico tomaría la forma de puntos aislados, idealmente como reactivos de opción múltiple para permitir economía en la calificación. Después de aplicación(es) piloto(s) y análisis de resultados en términos de la norma del grupo, y comparación con otras medidas (otras secciones del examen, y la opinión externa de un maestro que conociera a los examinandos), se establecerían calificaciones como criterios a diferentes niveles.

Se buscaría demostrar la validez concurrente de esta sección con uno o más instrumentos que pretendan evaluar conocimientos gramaticales. Existiría la posibilidad de utilizar algún instrumento ya disponible, o una selección de reactivos de exámenes ya existentes, para economizar en la elaboración del examen, siempre y cuando un estudio de los reactivos demostrara que lo que examinan está en conformidad con los objetivos de esta parte.

La sección integrativo-pragmática se basa en las investigaciones de Oller y otros que han demostrado la alta correla-

ción entre pruebas de dictado y cloze y otras pruebas, como de comprensión de lectura, producción escrita y comprensión auditiva. Aquí se incluyen pruebas de cloze y/o dictado, preparándolas como indica Oller.<sup>3</sup> Aquí se mide una combinación integrada de conocimientos gramaticales, de lexis, y del contexto lingüístico general desarrollado a través del texto utilizado. Se basa la validez de esta sección del examen en 1) la construcción "gramática de expectativas", esta combinación integrada de conocimientos internalizados que postula Oller, y 2) la validez concurrente con otras partes del examen, y con una opinión de un maestro externo que conoce a los examinandos, en aplicaciones piloto.

## 5.2 Evaluación del nivel comunicativo

La evaluación del nivel comunicativo también incluye dos acercamientos al problema. Por una parte, se evalúa el grado en que el examinando domina una serie de micro-habilidades que teóricamente son aplicables a una gran variedad de problemas y tareas no examinados. Por otra parte se incluyen tareas globales que miden la actuación comunicativa del candidato en una situación específica. Estos dos aspectos del nivel comunicativo no necesariamente se evalúan en dos sub-pruebas aisladas. Una sección sobre comprensión de conferencias, por ejemplo, podría incluir, intercaladas, tareas globales y la evaluación de micro-habilidades particulares. Sin embargo, el contenido y la construcción que miden, y el tipo de validación que requieren difieren suficientemente para merecer tratamientos independientes aquí.

### 5.2.1 Micro-habilidades

Esta sección refleja un análisis de la comunicación en una serie de micro-habilidades que subyacen a la actuación. Se espera en esta sección (con el examen de dominio lingüístico) obtener un instrumento que arroje información generalizable a tareas no evaluadas en la prueba.

Para seleccionar de la lista de Munby, un número de micro-habilidades a examinar, partimos de la lista de los eventos y actividades comunicativos arrojada por el análisis de necesidades, presentada en la sección 3.3 del presente trabajo. Se agruparon los eventos para hacer cuatro fases para el examen y luego se relacionaron, por medio de la introspección, cada actividad con las micro-habilidades que parecen contribuir a cada actividad. Este proceso es subjetivo, y no podemos asegurar que, para realizar una actividad dada, todo el mundo utilizaría todas las mismas micro-habilidades. Sin embargo, para una actividad dada, algunas micro-habilidades entran necesariamente, y es el único método que tenemos ahora para la selección.

La evaluación de estas micro-habilidades se basaría en una serie de tareas pequeñas diseñadas para forzar el uso de una de ellas.\* Se haría lo posible para diseñar estas tareas de modo que su calificación fuera objetiva, utilizando reactivos de opción múltiple, o falso-verdadero. Sin embargo, para

---

\* Hay ejemplos de estas tareas en la sección 6.2.

las secciones de producción oral y escrita, tendría que haber una calificación subjetiva. El análisis de los resultados sería parecido al que se describió ya para las secciones del nivel lingüístico en que un análisis en relación a la norma de un grupo piloto precedería al establecimiento de una calificación-criterio.

La validez de este tipo de subprueba se basa por una parte en el contenido a medir. Se haría la selección de micro-habilidades de manera de incluir las necesarias para la realización de las actividades comunicativas arrojadas por el análisis de necesidades. Por otra parte, la validez de construcción de la prueba depende de la validez de la teoría de una competencia comunicativa que incluye micro-habilidades. Sería posible, y muy interesante (aunque fuera del alcance del presente trabajo) intentar la validación de esta teoría a través de un estudio de tipo convergente-discriminante de dos o tres micro-habilidades.

#### 5.5.2 Tareas comunicativas globales

La tarea global pretende ser una medida directa, es decir, trata no de reflejar, sino de ser la misma cosa que lo que el candidato tendrá que hacer en el extranjero, en su programa académico. No cabe dentro de las actividades de un estudiante de posgrado escuchar un diálogo, y contestar preguntas de opción múltiple mientras tanto; no se incluye este tipo de actividad como tarea comunicativa global. Pero sí tiene que escuchar conferencias, tomar apuntes, y referirse

a ellos más adelante, para realizar otras tareas, como podría ser resumir la información para incluirla en un ensayo. También tiene que exponer temas a compañeros, y resolver sus dudas en seminarios. La inclusión de este tipo de tareas permite ver directamente la facilidad con que el candidato se comunica en situaciones representativas.\*

La evaluación del éxito del candidato en estas tareas se basa en criterios de dos tipos: 1) el éxito en la comunicación (¿aclaró la duda del compañero o no?) y 2) con relación a una serie de parámetros que permiten describir el nivel en que se comunica. La evaluación de esta parte tendría que ser subjetiva; se tendría que hacer un juicio deliberado acerca de la actuación de cada examinando, basado en los criterios mencionados.

La validez de esta sección por una parte es de contenido; las actividades medidas se obtienen de la lista arrojada por el análisis de necesidades. Y por otra parte, es visual -

---

\* Algunas de las "micro-habilidades" de Munby de hecho corresponden a tareas comunicativas. La micro-habilidad 43, "reducción del texto a través del rechazo de información redundante o irrelevante" corresponde a la toma de apuntes. La 40, "extracción de los puntos relevantes para resumir el texto completo, una idea o un tema específico, la idea o el punto subyacente al texto" corresponde a la producción de un resumen. Aquí en este trabajo, se consideran estas "micro-habilidades" como tareas comunicativas globales, ya que implican el funcionamiento de muchas otras micro-habilidades y conocimientos de la lengua, de manera integrada.

como se acerca a una medida directa, intuitivamente parece dar una información válida.

### 5.3 Practicabilidad

Con el diseño presentado, esperamos haber satisfecho las consideraciones prácticas mencionadas al principio de este capítulo. El diseño es flexible, y puede adaptarse a diferentes situaciones; puede ser más o menos económico según los recursos disponibles. Sobre todo se puede variar el examen en cuanto al número de áreas de especialidad en que se da el examen.

Las diferentes partes que se corrigen de manera objetiva contribuyen a la economía del diseño. Según los recursos disponibles, se puede especializar la prueba, adaptándola a las necesidades particulares de los examinandos en mayor o menor grado. En todo caso, la parte lingüística podrá ser única.

En el nivel comunicativo, el grado de especialización de los textos de comprensión, y temas de producción dependerá de los recursos que haya para la elaboración y aplicación de pruebas especializadas.

La evaluación de los resultados en términos de niveles de actuación (aquí son tres), facilita la interpretación de los resultados. La validez predictiva del diseño sólo, se podrá ver en la práctica, y con estudios de seguimiento de candidatos.

#### 5.4 El diseño del nivel comunicativo

En esta parte, se darán detalles acerca del diseño del instrumento de evaluación del nivel comunicativo del examen de requisito para becarios. No se entrará en más detalle en la evaluación de nivel lingüístico, ya que es una área en que más trabajo se ha hecho, que es más familiar para todos.

Aquí no se presenta un examen completo, sino los elementos necesarios para construir una serie de exámenes; en el capítulo VI se presentan unos ejemplos de reactivos y tareas para dar una idea concreta de la forma que podría tomar un examen hecho en base a los cimientos aquí presentados.

##### 5.4.1 Eventos y actividades comunicativos

La base para el diseño del instrumento para evaluar el nivel comunicativo es la lista de eventos y actividades comunicativos arrojada por el análisis de necesidades. Esta lista se encuentra en la sección 3.3.

##### 5.4.2 Los parámetros de evaluación y niveles de actuación

###### 5.4.2.1 Los parámetros de evaluación

Para evaluar de manera detallada el nivel en que se comunica el candidato (tanto a nivel productivo como receptivo), tomamos una serie de parámetros o características del texto y la actuación a evaluar. Para cada parámetro se establecieron niveles específicos que el examen pretende discriminar. Este sistema de establecimiento de niveles de actuación

lo maneja Carroll,<sup>4</sup> y es una adaptación de los "niveles meta" del modelo del perfil de necesidades de Munby.<sup>5</sup>

En su modelo, Munby considera dos aspectos que influyen en el nivel meta: las dimensiones del texto (incluyendo extensión, complejidad, rango, velocidad, flexibilidad y finura), y las condiciones de tolerancia (a error lingüístico, error estilístico, referencia al diccionario, interlocutor, etc., repetición, y vacilación). En el sistema de Munby, se asignan valores numéricos a cada uno de los parámetros, para los medios oral y escrito en forma receptiva y productiva. Entonces, se producen tablas de este tipo, para establecer niveles meta:

Dimensiones	Hablado		Escrito	
	Receptivo	Productivo	Receptivo	Productivo
Extensión				
Complejidad				
Rango				
Finura				
Velocidad				
Flexibilidad				
Condiciones de Tolerancia				
Error lingüístico				
Error estilístico				
Referencia				
Repetición				
Vacilación				

Para cada dimensión, y cada condición de tolerancia, y en cada modo, se asigna un valor del 1 (muy bajo/muy pequeño) a 7 (muy alto/muy grande).

#### CUADRO 5-2

PARAMETROS DE EVALUACION Y NIVEL META DEL MODELO DE MUNBY<sup>6</sup>

El sistema de parámetros de Carroll, muy parecido al de Munby, es el que se ha adoptado para el presente trabajo. En lugar de asignar valores numéricos para describir el nivel requerido para cada parámetro, Carroll presenta una descripción de cada nivel. Por ejemplo, en lugar de asignar un valor "1" a extensión en la columna "escrito-receptivo", se da una descripción de los textos. Podría decir, "textos desde una sola palabra hasta frases cortas." Esta manera de manejar los niveles se ha adoptado, porque nos parece más fácil de comprender, y de utilizar para basar decisiones acerca de cuáles textos incluir en un examen.

Carroll no distingue en sus parámetros entre los que se refieren al texto, y los que se refieren a condiciones de tolerancia, aunque se incluyen dimensiones de los dos tipos. Aquí resumimos los parámetros de Carroll, empleados en el presente trabajo:

Extensión	del texto
Complejidad	del texto
Rango	de funciones, tonos e intenciones en el texto.
Velocidad	con que maneja el texto, o de su producción
Flexibilidad	del candidato ante variedades de texto, hablantes, interlocutores, etc.
Adecuación	de los textos producidos por el candidato a la situación, y su capacidad de reconocer diferentes estilos y tonos.
Corrección	con la que produce sus mensajes, y con que comprende información tanto implícita como explícita.
Independencia	del candidato de fuentes de referencia (diccionario, interlocutor, etc.)
Vacilación	aparente a la hora de iniciar una comunicación
Repetición	que necesita el candidato para comprender un texto, o para hacerse entender

Los términos para identificar los parámetros son de Carroll, las explicaciones que siguen los términos se inferieron de las especificaciones de los parámetros a diferentes niveles, en los ejemplos que incluye Carroll. Para el diseño original, hemos incluido todos los parámetros que Carroll maneja; una versión piloto del examen contemplaría la simplificación del sistema, combinando algunos de los parámetros, y eliminando otros, de acuerdo con la tarea a evaluar.

#### 5.4.2.2 La especificación de los tres niveles de actuación para las cuatro fases del examen

El próximo paso en el diseño del examen fue establecer los niveles de actuación que se querían medir, en cada fase del examen. Para hacer esto, se hizo una definición clara y precisa de cada nivel en relación con los diez parámetros sugeridos por Carroll. Por lo pronto, se han establecido tres niveles de actuación que distinguiría el examen. El nivel 1, el más bajo, representa ya cierto manejo del idioma en tareas comunicativas; de ninguna manera representa el nivel de un principiante. El segundo nivel representa al individuo que, aparte de un manejo de la lengua en que apoyarse para realizar tareas comunicativas, también maneja una variedad de estrategias que le permiten comunicarse mejor. El último nivel, nivel tres, representa al individuo bastante fluido y sofisticado en su uso de la lengua, que puede realizar tareas comunicativas complejas, manejando textos extensivos.

Esta división en tres niveles es exploratoria, y antes de tener el examen completo y probado, sería difícil establecer el significado de estos resultados para un examen de requisito. Tentativamente, parece que una persona con una buena actuación en tareas al nivel dos, en todas las áreas del examen podría considerarse lista para empezar un programa de estudios en inglés, tal vez con el apoyo de instrucción en inglés y habilidades de estudio, que podría ser paralela a un curso académico, y que no se volvería actividad principal del becario. El individuo con una actuación al nivel 1 necesitaría instrucción en inglés y habilidades de estudio antes de iniciar un programa académico. Estas observaciones acerca de los niveles de actuación deben tomarse como completamente tentativas, y en necesidad de comprobación antes de que se implementaran en cualquier sistema de evaluación para becarios. Un sistema definitivo probablemente tendría que ser más fino, con dos o tres niveles de actuación entre los extremos definidos ahora por los niveles uno y tres. Para establecer los otros niveles, sin embargo, sería necesario comprobar el funcionamiento de los niveles ya existentes.

#### 5.5 La especificación de escalas de evaluación para el nivel comunicativo

A continuación, se encuentran los cuadros que dan las escalas de evaluación para los tres niveles, para cada fase del examen.

CUADRO 5-3

ESPECIFICACION DE NIVELES DE ACTUACION - COMPRENSION AUDITIVA (CONFERENCIAS)

PARAMETRO	NIVEL 1	NIVEL 2	NIVEL 3
1 Extensión	Maneja textos de dos a tres minutos de duración.	Maneja textos de cinco a diez minutos de duración.	Maneja textos de más de diez minutos de duración.
2 Complejidad	Puede manejar un tema principal con algunos detalles de apoyo, dada una presentación bastante clara.	Puede manejar uno o dos temas con sus detalles de apoyo. Dada una presentación relativamente clara, puede manejar algunas divagaciones.	Puede manejar varios temas con sus detalles de apoyo, a pesar de divagaciones del conferencista, ruido e interferencias.
3 Rango	Comprende conferencias sobre todo con temas muy familiares, con un rango limitado de funciones.	Comprende conferencias con cierta variedad de tonos y funciones, y con una gama de temas.	Comprende conferencias que incluyen una variedad amplia de tonos y funciones, y diferentes intenciones, y una variedad de temas. Reconoce la actitud del autor a través del tono.
4 Velocidad	Necesita que el conferencista hable de manera medida, aunque no exageradamente lento.	Entiende la mayor parte cuando el conferencista habla a velocidad normal.	Entiende, aún cuando el conferencista habla a una velocidad normal.

CUADRO 5-3 (Continuación)  
 COMPRENSION AUDITIVA (CONFERENCIAS)

PARAMETRO	NIVEL 1	NIVEL 2	NIVEL 3
5 Flexibilidad	Comprende al conferencista mientras no hayan ni cambios bruscos en el tema, ni divagaciones por el conferencista.	Puede manejar unos cambios no muy marcados en el tema y el tono del conferencista.	Puede manejar cambios de tema y tono en el conferencista, maneja divagaciones, y reconoce cuando éstas comienzan y terminan.
6 Adecuación a la Situación	No reconoce variaciones en el estilo, y tampoco se da cuenta de variaciones implicadas en chistes, ironía, etc.	Reconoce variaciones en el estilo del conferencista, mientras éstas no sean muy sutiles.	Reconoce y aprecia diferencias de estilos - conferencia formal, charla informal. Reconoce chistes, ironía, etc.
7 Corrección	Entiende de manera precisa la información explícita en la conferencia.	Entiende de manera precisa la información explícita en la conferencia.	Entiende de manera precisa la información explícita e implícita en la conferencia.
8 Independencia	Necesita referirse frecuentemente a material de apoyo (artículos, lectura citada) para recuperar información que perdió durante la conferencia.	A veces necesita referirse a material de apoyo para recuperar información que perdió durante la conferencia. Puede identificar sus lagunas.	Rara vez necesita referirse a material de apoyo para recuperar información que perdió durante la conferencia. Puede identificar sus lagunas.

CUADRO 5-3 (Continuación)  
 COMPRENSION AUDITIVA (CONFERENCIAS)

PARAMETRO	NIVEL 1	NIVEL 2	NIVEL 3
9 Repetición	Necesita frecuentemente pedir repetición de partes de la conferencia, ya sea al conferencista o a un compañero.	A veces necesita pedir clarificación de un punto; puede identificar los elementos que le hacen falta.	Rara vez necesita pedir clarificación o repetición de un punto; puede identificar los elementos que le hacen falta.
10 Vacilación	No aplicable.	No aplicable.	No aplicable.

CUADRO 5-4

ESPECIFICACION DE NIVELES DE ACTUACION - COMPRENSION DE LECTURA

PARAMETRO	NIVEL 1	NIVEL 2	NIVEL 3
1 Extensión	Maneja textos de dos a tres párrafos.	Maneja textos de dos a tres cuartillas.	Maneja textos de más de tres cuartillas.
2 Complejidad	Puede manejar un tema principal con algunos detalles de apoyo, dada una presentación bastante clara.	Puede manejar uno o dos temas principales con sus detalles de apoyo. Dada una presentación relativamente directa, puede manejar algunas divagaciones.	Puede manejar varios temas con sus detalles de apoyo, y las relaciones entre los temas.
3 Rango	Maneja textos sobre todo con temas familiares, limitados sobre todo al área de especialidad académica.	Comprende textos que tienen una variedad de temas y estilos.	Comprende textos con una amplia variedad de temas y tonos, y con diferentes intenciones. Reconoce el actitud del autor por medio del tono.
4 Velocidad	Lee muy detenidamente. Tiene buena comprensión cuando no tiene necesidad de leer rápidamente.	No lee muy despacio, palabra por palabra, sino que tiene cierta velocidad. Le faltan estrategias para revisar textos largos en búsqueda de información específica.	Lee rápidamente; puede revisar grandes cantidades de material en búsqueda de información específica. Puede captar bastante información con una lectura rápida.

CUADRO 5-4 (Continuación)  
 COMPRENSION DE LECTURA

PARAMETRO	NIVEL 1	NIVEL 2	NIVEL 3
5 Flexibilidad	Tiene dificultad en manejar más de una fuente para llegar a la solución de un problema.	Maneja información de dos o tres fuentes en la solución de un problema.	Maneja sin problemas una variedad de tipos de texto en la solución de un problema.
6 Adecuación a la Situación	Maneja sobre todo textos de su área de especialidad, en estilo directo.	Reconoce y aprecia los estilos más frecuentes para su situación: revista académica, revista de difusión.	Reconoce y aprecia diferencias de estilos. Reconoce chistes, ironía, etc.
7 Corrección	Entiende de manera precisa la información explícita.	Entiende de manera precisa la información explícita e implícita.	Entiende de manera precisa la información explícita e implícita.
8 Independencia	Necesita referirse con frecuencia al diccionario por falta de estrategias para comprender lexis nuevo, y por lo limitado de su vocabulario.	Solo necesita consultar el diccionario ocasionalmente; aunque tiene estrategias para comprender lexis nuevo, a veces comprensión limitada del contexto hace necesaria la consulta.	Rara vez necesita consultar el diccionario. Tiene vocabulario amplio, y estrategias para comprender lexis nuevo.

CUADRO 5-4 (Continuación)  
 COMPRENSION DE LECTURA

PARAMETRO	NIVEL 1	NIVEL 2	NIVEL 3
9 Repetición	Necesita re-leer el texto varias veces para encontrar y captar la información deseada.	Puede ser necesario re-leer partes del texto una o más veces para encontrar y captar la información deseada.	Generalmente encuentra y capta la información con una lectura, aunque en ocasiones puede ser necesario re-leer partes del texto para captar información detallada.
10 Vacilación	Puede tardarse en comprender la tarea; no tiene plena seguridad de lo que debe hacer para resolver el problema.	No necesita mucho tiempo para iniciar la tarea, aunque toma un momento para reflexionar sobre la tarea para organizar su acercamiento.	Inicia la tarea sin esperarse mucho; organiza rápidamente su acercamiento al problema.

CUADRO 5-5

ESPECIFICACION DE NIVELES DE ACTUACION - PRODUCCION ESCRITA

PARAMETRO	NIVEL 1	NIVEL 2	NIVEL 3
1 Extensión	Produce textos de uno o hasta dos párrafos.	Produce textos de dos a tres párrafos.	Produce textos de cinco y más párrafos.
2 Complejidad	Produce textos de organización sencilla, incluyendo sólo un mínimo de elementos de cohesión y coherencia.	Produce textos que incluyen la buena organización de varios puntos. Maneja bien varios elementos básicos de cohesión y coherencia.	Produce textos que incluyen varios puntos, bien organizados. Tiene un manejo de una gama amplia de elementos de cohesión y coherencia.
3 Rango	Sólo puede redactar sobre temas cotidianos, con un estilo directo.	Puede redactar sobre temas cotidianos, y, de manera limitada, de su área de especialidad académica.	Puede producir textos con una variedad amplia de temas y tonos.
4 Velocidad	No termina las tareas, o sólo produce textos muy cortos, en el tiempo dado, debido a su falta de velocidad a la hora de organizar y expresar sus ideas.	La velocidad con que <u>or</u> ganiza y expresa sus ideas le permite producir textos completos en el tiempo dado.	La velocidad con que organiza y expresa sus ideas le permite producir textos completos y extensivos en el tiempo dado pa-hacer las tareas.

CUADRO 5-5 (Continuación)

PRODUCCION ESCRITA

PARAMETRO	NIVEL 1	NIVEL 2	NIVEL 3
5 Flexibilidad	Maneja un número mínimo de funciones en sus textos, limitándose casi exclusivamente a la descripción y la narración.	Produce textos con un número mayor de funciones, como son: definir, clasificar, concluir.	Produce textos con una variedad de funciones, como son: definir, clasificar, opinar, resumir, concluir, comentar, etc.
6 Adecuación a la Situación	El estilo no siempre es apropiado para un examen escrito, ya que contiene muchas expresiones propias del lenguaje oral, o de situaciones informales.	Aunque su producción puede contener fallas estilísticas ocasionales, básicamente escribe un estilo apropiado a un examen escrito.	Produce un texto adecuado para un examen escrito, con un estilo formal, sin ser ni muy familiar ni muy artificial. Uso adecuado de expresiones idiomáticas.
7 Corrección	Tiene un manejo básico de los sistemas de gramática y lexis, a pesar de errores en el uso del sistema verbal, y en el orden de palabras, por ejemplo. Tiene errores más frecuentes con áreas problemáticas como preposiciones y artículos definidos e indefinidos.	Tiene un buen manejo de gramática y lexis, a pesar de errores mencionados en el nivel 1, pero menos frecuentes.	Alcanza un alto nivel de corrección léxico y gramatical. Maneja con relativa facilidad los sistemas verbales, y la estructuración de la oración. Puede tener problemas ocasionales con preposiciones y artículos definidos e indefinidos.

CUADRO 5-5 (Continuación)

PRODUCCION ESCRITA

PARAMETRO	NIVEL 1	NIVEL 2	NIVEL 3
8 Independencia	Se apoya frecuentemente en el diccionario, aunque no sabe hacer un buen uso de éste.	Se apoya frecuentemente en el diccionario, y sabe utilizarlo de manera eficaz.	Se apoya ocasionalmente en el diccionario, y sabe utilizarlo de manera eficaz.
9 Repetición	Hay mucha repetición de un número reducido de palabras y frases fijas.	Un vocabulario suficiente y un buen manejo de la lengua permiten un uso relativamente variado de la lengua.	Su amplio vocabulario y buen manejo de la lengua permiten un uso variado, no repetitivo de la lengua.
10 Vacilación	La falta de fluidez en la expresión escrita se refleja en la imposibilidad de terminar la tarea en el tiempo dado.	La relativa fluidez se refleja en el hecho que logra cumplir la tarea en el tiempo dado.	La fluidez en la expresión escrita se refleja en la extensión de los textos producidos en el tiempo dado.

CUADRO 5-6

ESPECIFICACION DE NIVELES DE ACTUACION -INTERACCION ORAL (SEMINARIO)

PARAMETRO	NIVEL 1	NIVEL 2	NIVEL 3
1 Extensión	Puede participar en una <u>in</u> teracción no muy prolongada, aunque pierde detalles. Participa de manera apropiada, aunque sus <u>in</u> tervenciones generalmente son breves, de unas palabras.	Puede participar en una interacción prolongada, con buena comprensión y participación, aunque a veces se le escapan detalles. Sus <u>in</u> tervenciones son a veces de cierta extensión.	Puede participar en una interacción prolongada, comprendiendo, reaccionando a y produciendo intervenciones extensivas, frecuentemente de varios enunciados.
2 Complejidad	Generalmente las <u>in</u> tervenciones incluyen un sólo punto; cuando son más extensivos, demuestran poca organización interna. Lingüísticamente, son muy sencillas.	Las <u>in</u> tervenciones demuestran cierta organización de información. Lingüísticamente, demuestra menos estrategias para evitar la complejidad.	Las <u>in</u> tervenciones a veces demuestran una organización de punto principal con detalles o ejemplos; incluyen comentarios, opiniones, etc.
3 Rango	Las <u>in</u> tervenciones demuestran un número mínimo de funciones. Expone lo preparado y tal vez pide aclaración.	Las <u>in</u> tervenciones incluyen cierta variedad de funciones.	Las <u>in</u> tervenciones incluyen una gama de <u>fun</u> cciones: pregunta, afirma, critica, opina, etc.
4 Velocidad	La producción fluye, aunque es lenta; hay pausas frecuentes. (Véase 10 Vacilación.)	Habla a velocidad casi normal, tal vez un poco mesurado; hay pausas ocasionales.	Habla a velocidad normal; a veces hay pausas.

CUADRO 5-6 (Continuación)  
INTERACCION ORAL (SEMINARIO)

PARAMETRO	NIVEL 1	NIVEL 2	NIVEL 3
5 Flexibilidad	Se limita básicamente a presentar lo que ha preparado. Demuestra poca flexibilidad ante temas o aspectos no preparados.	Demuestra más confianza cuando el tema es uno que ha preparado o que es familiar; sin embargo puede participar en la interacción sobre temas no preparados. Puede contestar preguntas y aclarar dudas acerca de sus intervenciones.	Puede participar en charlas informales sobre temas no preparados, y en presentaciones más formales. Puede interrumpir su presentación para responder a preguntas y aclarar dudas, tomar en cuenta una opinión contraria, etc., sin perder la secuencia de su presentación.
6 Adecuación a la Situación	La producción básicamente es apropiada a la situación de seminario, a pesar de fallas ocasionales. A veces no responde a preguntas, críticas u opiniones contrarias de manera apropiada.	La producción básicamente es apropiada a la situación del seminario, a pesar de fallas ocasionales, como el mal uso de frases idiomáticas. Responde al tono de las intervenciones de otros participantes.	Adecúa sus intervenciones a la situación global (seminario), y el contexto inmediato (responde a preguntas, críticas, opiniones contrarias, etc.). Responde al tono de las intervenciones de los otros participantes.
7 Corrección	Errores lingüísticos (de sintaxis, uso de falsos cognados, etc.) son frecuentes, sin embargo no interrumpen la comunicación.	Errores lingüísticos son ocasionales.	Errores lingüísticos son poco frecuentes.

CUADRO 5-6 (Continuación)  
INTERACCION ORAL (SEMINARIO)

PARAMETRO	NIVEL 1	NIVEL 2	NIVEL 3
8 Independencia	Ocasionalmente necesita preguntar al interlocutor por una palabra desconocida, o usa una palabra en español, por falta de estrategias que le permiten salir del "callejon". En sus presentaciones, se apoya mucho en sus apuntes; se le dificulta hablar espontáneamente.	Rara vez necesita preguntar al interlocutor para encontrar las palabras que necesita. Se apoya en sus apuntes para hacer presentaciones preparadas, pero puede salir de ellos para exponer extemporáneamente.	No necesita preguntar al interlocutor para encontrar palabras que necesita; tiene estrategias para hacer paráfrasis, etc. Puede salir de la exposición preparada, y hablar extemporáneamente.
9 Repetición	A veces pierde trozos largos de la interacción y necesita pedir que se repita. No siempre puede identificar el punto que perdió. Utiliza frases y palabras repetidas.	Ocasionalmente necesita pedir que se repita; puede identificar el punto que perdió. Un vocabulario suficiente, y un buen manejo de la lengua permiten un uso relativamente variado y no repetitivo de la lengua.	Rara vez necesita pedir que se repita, puede identificar el punto que perdió. Su amplio vocabulario y buen manejo de la lengua permiten un uso variado y no repetitivo de la lengua.
10 Vacilación	Habla con cierta fluidez, la vacilación frecuentemente refleja su falta de estrategias para compensar sus faltas de conocimientos lingüísticos.	Mantiene la fluidez aunque a veces vacila cuando busca una palabra desconocida o reformula para evitar áreas problemáticas.	No vacila en sus intervenciones. Tiene bastante fluidez. Cuando pausa, es más por la necesidad de organizar sus ideas, o considerar un punto, que por falta de conocimientos lingüísticos.

5.6 La especificación de las micro-habilidades a medir, con apuntes acerca de cómo se podrían medir

Habiendo establecido los niveles de actuación a medir en el examen, se prepararon listas de las micro-habilidades que consideramos que figuran en las diferentes áreas del examen. Se ha comparado esta lista de micro-habilidades con los tres niveles de ejecución para cada área. Se ha indicado, para cada micro-habilidad, cuáles de los tres niveles implican el dominio de esta micro-habilidad. Esta información, resumida en los cuadros de micro-habilidades, es esencial para la preparación de exámenes que pretenden evaluar al examinando en términos del criterio de los niveles. A continuación, se encuentran las listas de micro-habilidades implicadas por las actividades incluidas en cada una de las cuatro fases del examen. Se incluyen para cada micro-habilidad, apuntes tentativos acerca de cómo podría evaluarse.

CUADRO 5-7  
MICRO-HABILIDADES A MEDIR - COMPRENSION AUDITIVA (CONFERENCIAS)

Número de la habilidad *	Descripción	Niveles en que es necesaria			Cómo medir
		1	2	3	
1	Discriminación de sonidos en palabras aisladas.	x	x	x	Habilidad facilitadora básica; no se mide independientemente.
3	Discriminación de sonidos en la producción continua.	x	x	x	"
5	Discriminación de énfasis en la producción continua.	x	x	x	"
9	Reconocimiento del uso de énfasis en la producción continua.	x	x	x	"
13 15	Comprensión de patrones entonacionales; interpretación del actitud del hablante por medio de la variación de tono, cambio del núcleo, tesitura, pausa y velocidad.		x	x	Preguntas acerca de la actitud del hablante, cuando se muestra de esta manera. Aquí el examinando se apoya también en lexis, etc.
20	Comprensión de información explícitamente incluida en el texto.	x	x	x	Tarea global. Preguntas acerca del contenido.
22	Comprensión de información implícita, a través de inferencias, y lenguaje figurativo.			x	Tarea global. Preguntas acerca del contenido implícito.

\*El número de la habilidad se refiere a la taxonomía de Munby. <sup>7</sup>

CUADRO 5-7 (Continuación)  
 COMPRENSION AUDITIVA (CONFERENCIAS)

Número de la habilidad	Descripción	Niveles en que es necesaria			Cómo medir
		1	2	3	
26	Comprensión del valor comunicativo (función) de oraciones y enunciados.		x	x	Preguntas acerca de la función de partes del texto.
30 32	Comprensión de relaciones entre las partes de un texto por medio de la cohesión léxico - gramatical.	x	x	x	Preguntas directas acerca del contenido del texto que implican comprensión de los elementos de cohesión.
35	Reconocimiento de indicadores en el discurso.		x	x	Preguntas directas acerca de la función de diferentes partes de un texto. Interrumpir el texto después de un marcador de discurso, dando tres o cuatro posibles terminaciones, una de las cuales es congruente con el indicador.
37	Identificación del punto principal o información importante de un trozo de discurso.	x	x	x	Pedir que el examinando identifique la idea principal, o sea de manera abierta, o escogiendo de varias opciones.

CUADRO 5-7 (Continuación)  
 COMPRENSION AUDITIVA (CONFERENCIAS)

Número de la habilidad	Descripción	Niveles en que es necesaria			Cómo medir
		1	2	3	
39	Distinguir la idea principal de los detalles de apoyo.		x	x	Tareas que requieren la formulación de la estructura de la información del texto.
40	Extracción de los puntos relevantes para resumir el texto completo, una idea o un tema específico o la idea subyacente al texto.		x	x	Hacer que los examinandos tomen apuntes guiados del texto, y luego resumir la conferencia.
43	Reducción del texto a través del rechazo de información redundante o irrelevante.			x	Una tarea que requiere que el examinando tome apuntes, y utilizarlos después.
51	Transferencia de información de un código a otro.			x	Tareas que requieren que el examen crea o complete un diagrama, basándose en información presentada oralmente.

CUADRO 5-8  
MICRO-HABILIDADES A MEDIR - COMPRENSION DE LECTURA

Número de la habilidad	Descripción	Niveles en que es necesaria			Cómo medir
		1	2	3	
17	Reconocimiento de la escritura de la lengua.	x	x	x	Habilidad facilitadora básica; no se mide independientemente.
19	Deducción del significado de léxico desconocido.		x	x	Usar pseudo-palabras, en contexto. El examinando escoje entre tres o cuatro definiciones "equivalentes" en español.
20	Comprensión de información explícitamente expuesta.	x	x	x	Tarea global. Preguntas directas acerca del contenido.
22	Comprensión de información implícita, a través de inferencias y lenguaje figurativo.		x	x	Tarea global. preguntas directas acerca del contenido implícito.
24	Comprensión del significado conceptual, sobre todo: cantidad, definitividad e indefinitividad, comparación y grado, tiempo y aspecto, localización y dirección, medios, instrumento, causa, resultado, propósito, motivación, condición, contraste.		x	x	Preguntas acerca del contenido del texto que implican comprensión de estos conceptos.

CUADRO 5-8 (Continuación)  
 COMPRENSION DE LECTURA

Número de la habilidad	Descripción	Niveles en que es necesaria			Cómo medir
		1	2	3	
26	Comprensión del valor comunicativo (función) de enunciados y oraciones, con y sin indicadores explícitos.		x	x	Preguntas acerca de la función de partes del texto (véase micro-habilidad 35).
28	Comprensión de las relaciones dentro de la oración.	x	x	x	Medir en el nivel lingüístico del examen.
30	Comprensión de las relaciones entre las partes de un texto, por medio de estos tipos de cohesión lexical: repetición, sinónimos, hipónimos, antítesis, conjuntos léxicos y colocación.	(x)*	x	x	Preguntas acerca del contenido del texto que implican la comprensión de estas relaciones.
32	Comprensión de las relaciones entre las partes de un texto, por medio de estos tipos de cohesión gramatical: pro-formas, palabras generales, referencia, comparativos, sustitución, elipsis, marcadores de tiempo y espacio, nexos lógicos.	(x)	x	x	Preguntas acerca del contenido del texto que implican comprensión de estas relaciones.
34	Interpretación de un texto, sacando de ello. Referencia exofórica, leer entre líneas, integrar información en el texto con información previa.		x	x	Preguntas que piden que el examinando relacione el texto a sus conocimientos propios.

\*Habilidades indicadas entre paréntesis figuran en este nivel, pero en menor grado que en otros niveles.

CUADRO 5-8 (Continuación)  
 COMPRENSION DE LECTURA

Número de la habilidad	Descripción	Niveles en que es necesaria			Cómo medir
		1	2	3	
35	Reconocimiento de indicadores en el discurso para: introducir una idea, desarrollar una idea, transición de una a otra idea, concluir una idea, enfatizar un punto, explicar o clarificar un punto ya hecho, anticipación de una opinión contraria.		x	x	Preguntas acerca de la función de partes de un texto Interrumpir el texto después de un marcador de discurso, dando tres o cuatro posibilidades, una de las cuales es congruente con el indicador.
37	Indicar el punto principal de un trozo de discurso.	(x)	x	x	Subrayar en el texto tres partes, pedir que el examinando indique cuál corresponde al punto principal. Pedir que el examinando subraye el punto principal.
39	Distinguir la idea principal de los detalles de apoyo, diferenciando: significado primario y secundario, el conjunto de sus partes, un proceso de sus etapas, categoría de su exponente, afirmación de ejemplo, hecho de opinión, una proposición de su argumento.		x	x	Tareas que requieren la formulación de la estructura de la información en el texto.

CUADRO 5-8 (Continuación)  
 COMPRENSION DE LECTURA

Número de la habilidad	Descripción	Niveles en que es necesaria			Cómo medir
		1	2	3	
40	Extracción de los puntos relevantes para resumir el texto completo, una idea o un tema específico, la idea o el punto subyacente al texto.	(x)		x	Pedir un resumen, ya sea libre o guiado, de un texto.
41	Extracción de los puntos relevantes de un texto, implicando: la coordinación de información relacionada, la organización de información contrastiva acerca de un punto, la tabulación de información para análisis comparativo y contrastivo.			x	Tareas que requieren que el examinando saque información específica de un texto extensivo, o de una serie de textos, para colacarla en cuadros.
43	Reducción del texto a través del rechazo de información redundante o irrelevante, sobre todo: sacando elementos gramaticales, eliminando repetición, digresiones, circumlocución, comienzos falsos; comprensión de oraciones o grupos de palabras, uso de abreviaciones y símbolos.	(x)		x	Una tarea que requiere de la toma de apuntes (ya sea libre o guiada) de una lectura, y la referencia después a estos apuntes para llevar a cabo otra tarea. Se evaluaría la actuación del examinando en la segunda tarea, y no en sus técnicas personales para tomar apuntes.

CUADRO 5-8 (Continuación)  
 COMPRENSION DE LECTURA

Número de la habilidad	Descripción	Niveles en que es necesaria			Cómo medir
		1	2	3	
44	Habilidades básicas de referencia: comprensión y uso de representación gráfica (encabezados, subencabezados, numeración, sangrías, negritas, notas de referencia, índice y contenidos.			x	Tareas que requieren la revisión de una gran cantidad de material en poco tiempo, que implica el uso de estas habilidades para localizar rápidamente la información.
45	Skimming para obtener una apreciación global del contenido de un texto y una impresión general de ello.			x	Preguntas globales acerca de un texto, con limitaciones de tiempo.
46	Scanning para localizar información específica requerida, sobre uno o más puntos.			x	Ver tarea mencionada en micro-habilidad 44.
51	Transferencia de información presentada en un código (por ejemplo diagramático) a otro (por ejemplo lingüístico)			x	Tareas que requieren que el examinando complete o interprete una gráfica.
*	Uso eficiente del diccionario monolingüe o bilingüe.		x	x	Tareas de uso de diccionario.

\*Micro-habilidad no incluida en la taxonomía de Munby.

CUADRO 5-9  
MICRO-HABILIDADES A MEDIR - PRODUCCION ESCRITA

Número de la habilidad	Descripción	Niveles en que es necesaria			Cómo medir	
		1	2	3		
18	Manipulación de la forma escrita de la lengua.	x	x	x	Habilidad facilitadora básica; no se mide independientemente.	
21	Expresar información explícitamente.	x	x	x	Se evalúan todas las micro-habilidades a partir de textos producidos por el examinando en respuesta a una o dos tareas específicas, correspondientes a tareas comunicativas globales. Estas tareas deben diseñarse de tal manera que es posible ver el mayor número de micro-habilidades posible. La guía para la evaluación de la producción escrita se basaría en parte en esta lista de micro-habilidades.	
25	Expresar significado conceptual, especialmente: cantidad, definitividad e indefinitividad, comparación y grado, tiempo y aspecto, localización y dirección, medios, instrumentos, causas, resultado, propósito, motivación, condición, contraste.	(x)	x	x		
27	Expresar la función de oraciones con y sin indicadores explícitos.		x	x		
29	Expresar relaciones en la oración utilizando sobre todo elementos de la estructura de la oración, negación, modales y nexos.	x	x	x		
						"
31	Expresar relaciones entre las partes de un texto utilizando los elementos de cohesión lexical: repetición, sinónimos, pro-formas.		(x)	x		"

CUADRO 5-9 (Continuado)  
 PRODUCCION ESCRITA

Número de la habilidad	Descripción	Niveles en que es necesaria			Cómo medir
		1	2	3	
33	Expresar relaciones entre las partes de un texto utilizando los elementos de cohesión gramatical: referencia, comparativos, sustitución, elipsis, marcadores de tiempo y espacio, nexos lógicos.	(x)	x		Para la medición de estos aspectos, véase la hoja anterior.
36	Utilizar indicadores en el discurso para: introducir una idea, desarrollar una idea, hacer una transición de una idea a otra, enfatizar un punto, explicar o clarificar un punto ya hecho.	(x)	x		"
38	Indicar el punto principal o información importante en un trozo de discurso, por medio de oración de tema.	x	x		"
50	Planear y organizar información en lenguaje expositivo, utilizando funciones retóricas, como definición, clasificación, descripción, y procesos.	(x)	x		

CUADRO 5-10  
MICRO-HABILIDADES A MEDIR - INTERACCION ORAL (SEMINARIO)

Número de la habilidad	Descripción	Niveles en que es necesaria			Cómo medir
		1	2	3	
2	Articulación de sonidos en palabras aisladas	x	x	x	Habilidad facilitadora básica; no se mide independientemente.
4	Articulación de sonidos en la producción continua.	x	x	x	"
6	Articulación de patrones de énfasis en palabras.	x	x	x	"
8	Manipulación de variación de énfasis en la producción continua.	x	x	x	"
10	Manipulación del uso de énfasis en la producción continua.		x	x	"
12	Producción de patrones entonacionales con núcleo en posición neutra.		x	x	Se miden todas las micro-habilidades por medio de la observación de la actuación del individuo en una interacción tipo seminario. Esta interacción debe diseñarse tomando en cuenta los aspectos de la situación comunicativa enumerados por Morrow, y de modo que es posible ver el mayor número de micro-habilidades. La guía para la evaluación de la interac-
14	Producción de patrones entonacionales: expresión del significado de actitud, utilizando variación del tono, o cambio del núcleo.		x	x	
16	Expresión de significado de actitud, por medio de tésitura, pausa y ritmo.		x	x	

CUADRO 5-10 (Continuado)  
INTERACCION ORAL (SEMINARIO)

Número de la habilidad	Descripción	Niveles en que es necesaria			Cómo medir
		1	2	3	
21	Expresión de información explícitamente.	x	x	x	ción oral se basaría en parte en esta lista de micro-habilidades.
23	Expresión de información implícitamente, a través de inferencia y lenguaje figurativo.			x	
25	Expresión de significado conceptual, sobre todo: cantidad, definitividad e indefinitividad, comparación, y grado, tiempo y aspecto, localización y dirección, medios, instrumentos, causa, resultado, propósito, motivación, condición y contraste.	(x)	(x)	x	
27	Expresión del valor comunicativo (función) de enunciados, con y sin indicadores explícitos.		(x)	x	
36	Utilización de indicadores en el discurso para: introducir una idea, desarrollar una idea, hacer una transición de una idea a otra enfatizar un punto, explicar o clarificar un punto ya hecho.				

CUADRO 5-10 (Continuado)  
INTERACCION ORAL (SEMINARIO)

Número de la habilidad	Descripción	Niveles en que es necesaria			Cómo medir
		1	2	3	
47	Iniciación de un discurso: como iniciar, como introducir un punto nuevo, como introducir un tema nuevo.		x	x	Para la medición de estos aspectos, véase las hojas anteriores.
48	Mantener el discurso: como responder, seguir, adaptar a reacciones como tomar el turno, como ganar tiempo.		x	x	
49	Terminar el discurso, como salirse del discurso, como concluir un tema.		x	x	
	Algunas micro-habilidades relevantes del cuadro 5-7, Comprensión auditiva.				

## 5.7 Diseño de tareas de comprensión y de producción

Habiendo descrito los niveles a evaluar para cada parámetro, y enumerado las micro-habilidades, buscamos la manera en que se podría obtener acceso a información acerca del candidato en cada parámetro, y en relación a las micro-habilidades, de manera que se pudiera distinguir entre los tres niveles establecidos.

Carroll sugiere dos maneras diferentes de acercarse al problema de elaborar tareas para distinguir niveles de actuación.<sup>8</sup>

Por una parte, se puede presentar al candidato una serie de tareas de dificultad variable. Es decir, se incluyen algunas tareas fáciles que corresponden al nivel más bajo, algunas más difíciles para distinguir el próximo nivel, y así consecutivamente, hasta cubrir todos los niveles a evaluar en la prueba. Este acercamiento es especialmente apropiado para las habilidades receptivas, y lo hemos adoptado para las áreas de comprensión auditiva (conferencias) y comprensión de lectura. La graduación de las tareas se logra de dos maneras. Por una parte, para tareas que indican el nivel uno, sólo se incluyen micro-habilidades implicadas en este nivel, para el nivel dos, otros, etc. Pero por otra parte, los diferentes niveles implican también el manejo de distintos tipos de textos en cuanto a extensión, complejidad, etc. Entonces las tareas y los textos varían según el nivel que pretende discriminar esta parte del examen.

En el otro acercamiento que menciona Carroll, se pone una sola tarea, y se evalúa al examinando en cuanto al nivel que de muestra en la ejecución de ella. Este acercamiento se presta para evaluar la producción escrita, y la interacción oral.

### 5.7.1 Tareas de comprensión

Para crear una serie de tareas que permiten medir la comprensión, a tres diferentes niveles, como hemos dicho, se varían tanto los textos como las tareas relacionadas con micro-habilidades particulares.

Un instrumento piloto distinguiría la capacidad del examinando para manejar los diez parámetros a los tres niveles, y diferentes micro-habilidades. En el área de la comprensión algunos de los parámetros se refieren al texto, y otros a la actuación del candidato. Y otros se refieren a la actuación del examinando ante características específicas del texto. Para la comprensión de lectura, los parámetros se distribuyen de esta manera:

Parámetros que se refieren al texto	Parámetros que se refieren a la actuación del examinando ante características del texto	Parámetros que se refieren a la actuación del <u>examinando</u>
Extensión Complejidad Rango	Flexibilidad Adecuación a la situación	Velocidad Corrección Independencia Repetición Vacilación

CUADRO 5-11  
REFERENCIA DE LAS DIMENSIONES EN LA COMPRENSION DE LECTURA

Para la comprensión auditiva, la distribución es la misma, con excepción del parámetro "Velocidad". En la comprensión, auditiva, ésta es característica del texto, no de la actuación del examinando.

Los parámetros que se refieren al texto, analizados a diferentes niveles, forman la base para la selección de textos de comprensión para los tres niveles.\* Los parámetros de actuación, en combinación con la lista de micro-habilidades, guían la elaboración de tareas en relación con los textos seleccionados. De esta manera, se sistematiza el proceso de escribir pruebas de comprensión, basándose en criterios pre-establecidos. Así se asegura que el examen mide, en el grado en que sea posible, los elementos que se quieren medir, y no simplemente los elementos a que se presta un texto elegido al azar.

### 5.7.2 Tareas de producción

En las tareas de producción, todos los parámetros se refieren a la actuación del examinando. Sin embargo, podemos distinguir entre los parámetros más generales, y los que se refieren a una actuación específica ante un tipo particular de tarea. El cuadro 5-12 divide los parámetros en estos dos grupos. Los parámetros específicos sólo se podrán observar en

---

\*Véase Shaw, en prensa, para una manera eficiente de organizar la búsqueda de textos adecuados para trabajar, o evaluar, diferentes micro-habilidades.

la actuación del examinando, si se diseñan tareas que requieren una producción de tipo particular, mientras que los parámetros generales se podrán observar en prácticamente cualquier tarea de producción. El diseño de tareas comunicativas globales para la producción, entonces, tendrá que tomar muy en cuenta, sobre todo los parámetros del segundo grupo.

Parámetros generales	Parámetros específicos
Extensión	Complejidad
Velocidad	Rango
Independencia	Flexibilidad
Repetición	Adecuación a la
Vacilación	situación
Corrección	

CUADRO 5-12  
PARAMETROS GENERALES Y ESPECIFICOS

Notas de referencia bibliográfica

Capítulo V

<sup>1</sup> Helmut Vollmer, "Why are we interested in 'General Language Proficiency'?", en Issues in Language Testing, eds., J. Charles Alderson y Arthur Hughes (London: The British Council, 1981), pp. 170-171.

<sup>2</sup> J. Charles Alderson, "Report of the Discussion of General Language Proficiency," en Issues in Language Testing, eds. Alderson y Hughes.

<sup>3</sup> John W. Oller, Language Tests at School: A Pragmatic Approach (London: Longman, 1979), pp. 267-297; 363-397.

<sup>4</sup> Brandon J. Carroll, Testing Communicative Performance: An Interim Study (Oxford: Pergamon Press, 1980), p. 30.

<sup>5</sup> John Munby, Communicative Syllabus Design: A Sociolinguistic model for defining the content of purpose specific language programmes. (Cambridge: Cambridge University Press, 1978), pp. 91-97.

<sup>6</sup> Ibid.

<sup>7</sup> Ibid., pp. 123-131.

<sup>8</sup> Carroll, Testing Communicative Performance, p. 31.

## CAPITULO VI

### EJEMPLOS Y CONCLUSIONES

#### 6.0 Introducción

En este capítulo primero se presentan dos ejemplos de sub-pruebas del examen comunicativo en el área de comprensión de lectura. Luego se dan conclusiones generales del trabajo.

#### 6.1 Los ejemplos

Para ejemplificar el funcionamiento del diseño del nivel comunicativo del examen, incluimos dos sub-pruebas, diseñadas para evaluar la comprensión de lectura, al nivel dos (vease Cuadro 5- 5 Niveles de Actuación - Comprensión de Lectura).

Un ejercicio demuestra la evaluación de una serie de micro-habilidades. El otro, basado en el mismo texto, es una prueba de actuación, en el sentido que pide al examinando que haga una tarea lo más cercana posible a lo que tendrá que hacer como estudiante en el extranjero. En este caso, debe hacer un resumen de un texto de tres páginas. El resumen se evalúa de manera subjetiva, teniendo en cuenta una respuesta modelo que incluye los puntos considerados como esenciales. Esta tarea se acerca más a una medida directa, y balancea en el examen completo la prueba de micro-habilidades que presenta tareas bastante artificiales.

### 6.1.1 Introducción al texto

El texto: Gino Germani, "Definitions of Urbanization"

El texto "Definitions of Urbanization" es la primera sección del primer capítulo (Urbanization, Social Change and the Great Transformation) del libro Modernization, Urbanization and the Urban Crisis, editado por Germani. El libro incluye una serie de artículos importantes en el estudio de la urbanización, y el primer capítulo sirve de introducción a los artículos. Incluye definiciones y relaciona el trabajo de diferentes contribuciones al volumen. En la primera sección, habla de definiciones del término "urbanización".

Extensión: 3 páginas (diez párrafos, aproximadamente 110 líneas, aproximadamente mil doscientas palabras).

Complejidad: Establece dos maneras de definir la urbanización define y da ejemplos de cada uno de ellos. Trata problemas con uno de los dos tipos de definiciones. Trata los dos tipos de definiciones en relación al trabajo de varios autores, y llega a una conclusión acerca de la definición del término.

Rango: Incluye definiciones, ejemplos, clasificación, descripción, conclusiones.

A continuación reproducimos el texto de Germani:

## Urbanization, Social Change, and the Great Transformation

### DEFINITIONS OF URBANIZATION

Not only are the terms "urbanization" and "modernization" frequently used in contemporary social science, but since they are thought to represent aspects of the "revolution of our time," these words have been taken up by politicians, journalists, and others in a variety of social contexts. Such diffusion, however, does not guarantee clarity or an unambiguous meaning. On the contrary, as so often happens with social science concepts, several definitions of them are possible and, in fact, are currently employed.

The two meanings of urbanization in which I am interested here are the *demographic* and the *sociological*. The former seems less complex and controversial than the latter. It requires only two criteria — space and population (that is, the size and density of a human settlement). "Urban" is defined in terms of an agglomeration of population of a given size, within a certain area. Demographic urbanization, then, proceeds in two ways: "The multiplication of the points of concentration and the increase in size of individual concentrations."<sup>1</sup> It must be remembered, however, that even this relatively simple statement generates different interpretations and applications. For instance, although official statistics and census reports usually adopt demographic definitions, they vary with regard to the dividing line (in terms of size) between urban and non-urban. A common procedure is to define as "urban" those places with 2000 or more inhabitants; but the numbers 5000, 10,000, or 20,000 have sometimes been used instead to define minimum size.

On occasion, demographers and official statisticians have adopted entirely different criteria to define "urban," such as the type of politico-administrative organization of an area or, as was suggested in the past by the International Institute of Statistics, the preponderance of agricultural or

This chapter was written while I was at the Center for Advanced Studies in the Behavioral Sciences, in Palo Alto, California. I would like to express my thanks to the Center for the help I received while I was there.

non-agricultural occupations among the inhabitants.<sup>2</sup> Economic, legal, or political criteria are clearly related to a sociological definition rather than to a purely demographic conception of urbanization. A sociological definition of "urban" should take social structure and psychosocial and behavioral patterns into account, postulating that urban society differs from non-urban society in some significant way.

To try to endow a definition of this type with universal validity — that is, to cover all possible types of urbanization (and cities) — may turn out to be incomplete and insufficient for many scientific purposes. The reason lies, of course, in the fact that the nature of the "urban" or the "non-urban" changes according to the type of society. These concepts, like many others in the social sciences, are best dealt with by taking into consideration specific historical conditions. Except perhaps in the case of very general and highly abstract theorizing, the study of urbanization requires a lower level of generality, more concrete concepts, definitions carefully fitted to actual historical and cultural conditions. Obviously, different problems may require diverse levels of generality and abstraction. The objective should be to choose the most appropriate level in each case and to avoid the extension of the application of a given concept beyond the sociohistorical area in relation to which it was constructed.

Louis Wirth's classic definition of the city as "a relatively large, dense, and permanent settlement of socially heterogeneous individuals" and his psychosocial analysis of "urbanism as a way of life" certainly constitute one of the best attempts to formulate a theoretical framework for the study of urbanization and the city.<sup>3</sup> Wirth incorporates the essentials of the existing sociological tradition on this subject, to which I will refer later. Volume, density, and heterogeneity may be considered the structural independent variables in the definition. The remaining traits, mostly psychosocial in nature, are dealt with as dependent variables characterizing *urbanism*: predominance of secondary over primary groups, individualism, tolerance, segmental roles, propensity for change, social and ecological mobility, abstract thinking, universalism, achievement, instrumentality and so forth, along with other traits regarded as the "dark" side of urbanism, such as anomie, social disorganization, loss of identity, and other aspects largely analyzed in the usual critique of "mass society"<sup>4</sup>

Notice that Wirth's analysis corresponds to a given historical phase of urban history and perhaps to a specific cultural tradition and socio-economic setting. Many of the psychosocial traits included are precisely those usually attributed to "industrial society" in the commonly used dichotomies (or *continua*) opposing "traditional" (or sacred) to "modern" (or secular) society. Wirth's formulation therefore cannot be regarded as a universally valid definition of urbanization. It pertains essentially to "modern" urbanization or even more particularly to one of its different types.<sup>5</sup> Analogous limitations and qualifications have been applied to other conceptions of the urban — such as the "folk-urban" continuum.

Two points deserve emphasis here. Methodological criticism of this sort is correct. Yet such criticism should not lead us to forget two well-known facts: that modern society is an *urban* society par excellence and that the city has played an essential role in the rise of modernity. The modernizing role of cities is one of the central themes of this book.

Typologies and distinctions are the usual available procedures to overcome the difficulties and limitations of universal definitions or the improper extension of historically or culturally limited ones. Max Weber's distinction between the occidental and the oriental city is crucial in the historical analysis of the rise of modernity in the West.<sup>6</sup> Similarly, the typology presented by Redfield and Singer, contrasting "folk" society, "primary" urbanization (and cities of "orthogenetic transformation") with "secondary" urbanization (and cities of "heterogenetic transformation"), not only answers some of the criticisms addressed to the "folk-urban" continuum construct but also provides an excellent basis for the analysis of the relation between urbanization (the city of *heterogenetic* transformation) and the process of secularization, one of the central processes of modernization.<sup>7</sup> An analogous analytic function is achieved by the three types of urbanization, identified by Pizzorno, in one of the selections included in this volume.<sup>8</sup>

In the same vein, the evolutionist typologies developed by several authors may be fruitful for analytic and descriptive purposes. Sjöberg's *Pre-Industrial City* and the "phases" of urbanization suggested by Eric Lampard may be cited as good illustrations.<sup>9</sup> "Primordial urbanization," according to Lampard, is defined as "the first achievement of incipient urban organization as an *additional* and more *productive* mode of collective adaptation to physical and social environment." Definitive urbanization for him is the "culmination of primordial tendencies in the additional and alternative forms of social organization." Because of its new capacities, the "definitive city artifact is now capable of transplanting itself from its native uterine environments." Definitive urbanization is then distinguished in two sub-stages, "classic urbanization" (pre-industrial civilization) and "industrial urbanization."

In this chapter and in the readings a variety of definitions will crop up. Perhaps the best solution for the study of urbanization would be the combination of demographic and sociological definitions, the former being constructed or selected according to the nature of the problem at hand.

#### SECULARIZATION, CIVILIZATION, MODERNIZATION

Definitions of modernization are much more controversial, and I will not discuss them. Instead I will use the notion of *secularization*, a concept associated both with the city and with the two great changes in the history of mankind: the transition from "primitive" to "civilized" society and the

6.1.2 La prueba de micro-habilidades

6.1.2.1 Los reactivos

Reactivo

Habilidad y elemento evaluado

1. Cuáles términos tienen varias definiciones según el autor? (Párrafo 1)

- a) Urbanización y modernización
- b) Revolución de nuestros tiempos
- c) a y b

2. En el párrafo dos se establecen criterios para:

- a) una definición sociológica
- b) una definición demográfica
- c) tanto una definición sociológica como demográfica

3. Complete el siguiente cuadro con información del texto.

Dos definiciones de urbanización:

1) \_\_\_\_\_ 2) \_\_\_\_\_

Criterios para establecer las definiciones:

a) \_\_\_\_\_ a) \_\_\_\_\_  
b) \_\_\_\_\_ b) \_\_\_\_\_  
c) \_\_\_\_\_ c) \_\_\_\_\_

32. Cohesión gramatical:  
These words, they

32. Cohesión gramatical:  
the former, it

20. Comprensión de información explícita. Requiere de la comprensión de un trozo de dos párrafos, y la localización de información específica.

4. Falso Verdadero Hay acuerdo general acerca de la definición demográfica de la urbanización.
5. El párrafo cuatro habla de los problemas con:
- a) definiciones sociológicas
  - b) definiciones demográficas
  - c) tanto definiciones sociológicas como demográficas.
6. El trabajo de Wirth incluye elementos de definiciones:
- a) demográficas
  - b) sociológicas
  - c) tanto sociológicas como demográficas.
7. Qué tienen en común los trabajos de Weber, Redfield, Singer, Rizzorno, Sjoberg y Lampard?
- a) Se hicieron con base en la observación concreta del proceso de urbanización
  - b) Cada uno de ellos identifica más de un tipo de urbanización
  - c) Todos tratan la evolución de las ciudades.
  - d) Todos dan una definición universal de la urbanización.
22. Comprensión de información implícita. El texto menciona "different interpretations and applications", dice que reportes estadísticos "varían", y cita cuatro definiciones que se usan con cierta frecuencia.
30. Cohesión léxico - gramatical:  
A definition of this type.
- 22 Comprensión de información implícita.
- 20 Comprensión de información explícita.
- 22 Comprensión de información implícita
- 39 Distinguir entre afirmación y ejemplo.

8. Germani incluye descripciones del trabajo de Weber, Redfield y Singer, Rizzorno, Sjoberg y Lampard como:

- a) etapas en el desarrollo de la definición de la urbanización
- b) ejemplos de definiciones apropiadas para propósitos particulares
- c) elementos de una definición global de la urbanización
- d) ejemplos de definiciones no adecuadas de la urbanización.

Ponga el número de la parte subrayada del texto, que corresponde a cada descripción.

9. \_\_\_\_\_ Un ejemplo para apoyar una generalización.

10. \_\_\_\_\_ El resumen del argumento.

11. \_\_\_\_\_ La introducción de un punto que quita fuerza a lo anterior.

12. \_\_\_\_\_ Dar otro ejemplo para apoyar a un mismo punto.

13. Es propósito de Germani en estas tres páginas es:

- a) explicar cuál es su definición de la urbanización.
- b) revisar las diferentes definiciones de la urbanización
- c) demostrar que la definición de la urbanización debe variar según el estudio del fenómeno.
- d) demostrar que las definiciones utilizadas hasta la fecha no son apropiadas.

26. Comprensión del valor comunicativo de un trozo de discurso.

26. Comprender el valor comunicativo - ejemplificar: for instance

- resumir

32. Comprender relaciones entre partes de un texto, por medio de nexos lógicos: however.

26. Comprender el valor comunicativo - ejemplificar, apoyar un punto ya hecho: in the same vein.

37. Identificar el punto principal de un trozo de discurso.

### 6.1.2.2 Análisis de los reactivos en términos de las micro-habilidades del nivel 2.

Aquí vemos una lista de todas las micro-habilidades del nivel dos, sacada del cuadro 5-5, indicando cuáles se evalúan con las trece reactivos sobre el texto de Germani.

Micro-habilidades relevantes Nivel 2	Número de reactivos que la mide
19. Deducción de léxico desconocido	0
20. Información explícita	2
22. Información implícita	3
24. Significado conceptual	0
26. Valor comunicativo	4
30. Cohesión léxica	1
32. Cohesión gramatical	3
34. Relación del texto con conoci- mientos previos	0
35. Indicadores de discurso	0
37. Indicar el punto princi- pal	1
39. Distinguir entre afirmación y ejemplo	1

Un análisis de este tipo permite ver precisamente lo que se ha incluido en una serie de reactivos, rápidamente. Esto permite refinar o ampliar el instrumento de manera que refleje mejor lo que se quería medir. Aquí, vemos que se ha cubierto la mayoría de las micro-habilidades a evaluar para el nivel 2, aunque sea con un número mínimo de reactivos (de uno a cuatro) para cada uno de ellas, con excepción de las habilidades 19, 24, 34 y 35, que no se incluyeron. Se debe notar que las habilidades 26, 30 y 35 todas se refieren a un número amplio de elementos: a diferentes funciones, varios tipos de cohesión léxico y gramatical, y diferentes indicadores de discurso. La habilidad 39 se refiere no sólo a distinguir entre afirmación y ejemplo, sino también entre la idea principal y deta-

lles de apoyo, el conjunto y sus partes, un proceso y sus etapas, etc. En estos casos, sólo hemos evaluado uno, dos o tres de estos posibles elementos. Este hecho, con la falta total de reactivos que se refieren a varias de las habilidades, apunta la necesidad de ampliar la prueba, incluyendo por lo menos otro texto de extensión parecida, con otra serie de reactivos, para medir micro-habilidades a nivel 2.

### 6.1.3 La prueba de actuación

#### 6.1.3.1 Instrucciones para la prueba de actuación

Resuma las definiciones de urbanización discutidas por Germani, y sus conclusiones acerca de la mejor definición para la investigación del fenómeno. El resumen puede ser en español o inglés, y puede tomar la forma de un párrafo o de apuntes breves.

#### 6.1.3.2 Comentarios sobre la prueba de actuación

Esta prueba de actuación, como ya se mencionó, trata de producir en el examen, una situación paralela a la que se enfrentará el examinando como estudiante en el extranjero. No es únicamente una prueba de lectura; entra una serie de capacidades no totalmente analizadas que permiten al candidato escribir un resumen. No pretende ser una prueba "limpia" que excluya factores ajenos a la lectura. De hecho, pide que el examinando ponga en juego todo un proceso, mal entendido por examinadores e investigadores, en la solución de un problema específico.

### 6.2 Comentarios sobre las dos pruebas

Estas dos pruebas de comprensión de lectura ejemplifican la evaluación de las dos partes del nivel comunicativo del

diseño del examen, en el área de comprensión de lectura. Demuestra la manera en que se ha tratado de diversificar la manera de evaluar al candidato, al nivel comunicativo. Así mismo se ha diversificado el examen completo, incluyendo un nivel lingüístico y un nivel comunicativo.

Esta diversificación responde a la falta de respuestas claras a preguntas como ¿qué significa conocer una lengua? y ¿cuál es la mejor forma de medir estos conocimientos? Dada esta falta, consideramos que la manera de proceder para obtener el instrumento que sea el más justo, válido y confiable posible, es hacer un examen que refleje la variedad de posibles respuestas a estas preguntas básicas, y que refleje las necesidades del examinando.

### 6.3 Conclusiones sobre la evaluación basada en necesidades

En el capítulo cinco, hemos presentado el diseño detallado de un examen para candidatos para becas a estudiar en países de habla inglesa. Para hacer este diseño hemos tenido en cuenta dos áreas de investigación. Por una parte, las necesidades comunicativas de los candidatos se tomaron en cuenta de manera sistemática en el diseño. Por otra parte, el diseño del examen completo se hizo tomando en cuenta algunas investigaciones sobre la evaluación.

Se presenta aquí un diseño que consideramos es balanceado, y

esperamos que podrá reflejar la capacidad del futuro becario para enfrentarse a las exigencias de estudiar en un país de habla inglesa. Después de analizar algunos estudios sobre la evaluación, llegamos a la conclusión que la investigación sobre la evaluación, y aún sobre la definición de dominio lingüístico y competencia comunicativa no indica un sólo camino claro para la elaboración de instrumentos de evaluación. Hay varias definiciones tentativas de lo que significa dominar una lengua, y tener competencia comunicativa para uno u otro propósito. Por lo tanto, vimos la necesidad de diseñar una prueba diversificada que tomaría en cuenta los diferentes tipos de capacidades que influyen en la habilidad de un individuo de enfrentarse a una situación comunicativa específica, y también a una variedad de situaciones.

En este sentido, el trabajo presenta una respuesta tentativa, todavía a probarse y modificarse, al problema de la evaluación de la competencia comunicativa para un propósito específico. También presenta un procedimiento posible para desarrollar exámenes basados en criterios establecidos en relación a las necesidades comunicativas de los candidatos. Mientras que los detalles del diseño presentado se refieren al examen de requisito para candidatos a becas, esperamos que sean de interés más general, para personas involucrados en el proceso de la evaluación.

## REFERENCIAS

- Alderson, J. Charles. Issues in Language Testing. London: The British Council, 1981.
- Allen, J. B. B. y Davies, Alan, eds. The Edinburgh Course in Applied Linguistics Volume 4 Testing and Experimental Methods. London: Oxford University Press, 1977.
- Allen, J. P. B. y Corder, S. Fitt, eds. The Edinburgh Course in Applied Linguistics Volume 3 Techniques in Applied Linguistics. London: Oxford University Press, 1974.
- Anderson, Jonathon. Psycholinguistic Experiments in Foreign Language Testing. St. Lucia, Queensland: University of Queensland Press, 1976.
- Bachman, L. F. y Palmer, A. S. "The Construct Validation of Tests of Communicative Competence." Trabajo presentado en la 16a Convención Anual de TESOL, Honolulu, Hawaii, 1-6 mayo, 1982.
- The British Council, Study Modes and Academic Development of Overseas Students. London: The British Council, 1980.
- Brumfit, C. J. y Johnson, K., eds. The Communicative Approach to Language Teaching. Oxford: Oxford University Press 1979.
- Canale, Michael y Swain, Merrill. "Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing." Applied Linguistics I,1.
- Carroll, Brandon J. Testing Communicative Performance: An Interim Study. Oxford: Pergamon Press, 1980.
- Colle, Gwenn; Hildreth, Ann; y Howard, David. 'Proyecto Semestral para Diseño de Cursos. CELE/UNAM, 1981.
- Davies, Alan, ed. Language Testing Symposium: A Psycholinguistic Approach. London: Oxford University Press, 1968.
- Davies, Alan. "Language Testing" Language Teaching and Linguistics Abstracts, XI, 3-4.
- Fox, Robert. American University, Washington, D.C. Entrevista, Primavera, 1981.

- Harris, David P. Testing English as a Second Language. New York: McGraw Hill, 1969.
- Howard, David y Hildreth, Ann. Reseña de Communicative Syllabus Design, de John Munby. Estudios de Linguística Aplicada, en prensa.
- Johnson, Keith. Communicative Syllabus Design and Methodology. Oxford: Pergamon Press, 1972.
- MacKay, Ronald y Mountford, Alan, eds. English for Specific Purposes. London: Longman, 1978.
- Munby, John. Communicative Syllabus Design. A sociolinguistic model for defining the content of purpose specific language programmes, Cambridge: Cambridge U. Press, 1978.
- Oller, John W. Language Tests at School: A Pragmatic Approach. London: Longman, 1979.
- Oller, John W. y Perkins, Kyle, eds. Language in Education: Testing the Tests. Rowley; Newbury House Publishers, 1978.
- Oller, John W. y Perkins, Kyle, eds. Research in Language Testing. Rowley: Newbury House Publishers, 1980.
- Oller, John W. y Richards, Jack C. eds. Focus on the Learner: Pragmatic Perspectives for the Language Learner. Rowley: Newbury House Publishers, 1973.
- Palmer, Adrian S.; Groot, Peter J. M.; y Trosper, George A., eds. The Construct Validation of Tests of Communicative Competence. Washington, D.C.: TESOL, 1981.
- Richterich, René; y Chancerel, Jean-Louis, Identifying the Needs of Adults Learning a Foreign Language. Oxford: Pergamon Press, Council of Europe, 1977.
- Schultz, Renate. "Discrete Point versus Simulated Communication Testing in Foreign Languages." The Modern Language Journal LXI, 3 (1977).
- Shaw, Anthony, "1.2/1 Diseño de Cursos." Manuscrito, Biblioteca del CELE, UNAM.
- Stevens, Peter. New Orientations in the Teaching of English. Oxford: Oxford University Press, 1977.
- Valette, Rebeca. Modern Language Texting, 2nd edition. New York: Harcourt Brace Jovanovich, 1977.