



Literatura y Lingüística

ISSN: 0716-5811

literaturalinguistica@ucsh.cl

Universidad Católica Silva Henríquez
Chile

Vine Jara, Ana; Ferreira Cabrera, Anita
Propuesta de un modelo para una prueba con fines específicos académicos en ele
Literatura y Lingüística, núm. 33, 2016, pp. 369-390
Universidad Católica Silva Henríquez
Santiago, Chile

Disponible en: <http://www.redalyc.org/articulo.oa?id=35245697018>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

PROUESTA DE UN MODELO PARA UNA PRUEBA CON FINES ESPECÍFICOS ACADÉMICOS EN ELE*

Ana Vine Jara**

Anita Ferreira Cabrera***

Resumen

En este artículo se propone un modelo para construir un prototipo de prueba de Español como Lengua Extranjera con Fines Académicos (EFA), en línea, para un nivel B2. Para ello, se realizó un estudio mixto de dos etapas: 1) Cualitativa: Se definió los criterios necesarios para diseñar un prototipo de prueba de EFA en línea de nivel B2 y 2) Cuantitativa: Se utilizaron estos datos para la elaboración y validación del instrumento por expertos. Los resultados indican que el prototipo de prueba es válido en cuanto evalúa las distintas habilidades lingüísticas del nivel B2 de EFA en un entorno interactivo y acorde con el perfil de los estudiantes.

Palabras claves: Evaluación, español con fines académicos, criterios, validación.

A PROPOSAL OF A MODELTO FOR A TEST OF SPANISH AS A FOREIGN LANGUAGE FOR SPECIFIC ACADEMIC PURPOSES

Abstract

This article suggests a model to construct a prototype of an online B2 level test of Spanish as a Foreign Language for Academic Purposes (SAP). To this end, we carried out a mixed study in two stages: 1) Qualitative: The necessary criteria for the design of an online test of SAP for level B2 were defined, and 2) Quantitative: The gathered data was used for the test construction and validation by experts. Results indicate that the test prototype is valid, as it evaluates the B2 related linguistic skills for SAP in an interactive environment and in agreement with students' profiles.

Key words: Evaluation, Spanish for academic purposes, criteria, validation.

Recibido: 25-08-2015

Aceptado: 04-11-2015

* El estudio que se presenta forma parte del proyecto Fondecyt 1140651: "El feedback correctivo escrito directo e indirecto en la adquisición y aprendizaje del español como lengua extranjera"

** Doctora en Lingüística. Universidad de Concepción.

*** Doctora en Lingüística. Universidad de Concepción.

Introducción

La evaluación de lenguas con fines específicos se diferencia de los fines generales porque considera el uso de la lengua en ámbitos especializados. Sin embargo, existen dudas respecto de lo que se considera general y específico en una prueba.

La Evaluación del Español como Lengua Extranjera (ELE) con fines académicos presenta una serie de problemáticas relacionadas con la falta de acuerdo entre los especialistas respecto de lo que se está midiendo en una prueba de Español como Lengua Extranjera con Fines Específicos Académicos (EFA) y la necesidad de contar con estudios que muestren cómo se han diseñado, desarrollado y validado las pruebas existentes (Llorián, 2012). A esto se suma que, actualmente, no hay pruebas de EFA que tengan un reconocimiento nacional o internacional (Mendoza, 2015), lo cual sitúa al español en una situación desfavorable en el panorama de la certificación lingüística internacional. Del mismo modo, se necesita avanzar hacia formatos actualizados de aplicación de pruebas, acorde con las nuevas realidades educativas (García y Navarro, 2008). En este estudio se presenta un modelo de criterios evaluativos para el diseño de un prototipo de prueba de EFA, en línea, de nivel B2 que permita evaluar a los estudiantes extranjeros de intercambio universitario. El formato de aplicación automatizado del instrumento representa una ventaja, pues permitirá evaluar a los estudiantes aún antes de llegar a la universidad, para así apoyarlos con cursos de español con fines académicos a quienes carezcan de estas competencias (Ferreira, Vine y Elejalde, 2013).

La evaluación del EFA es una necesidad, tanto a nivel nacional como internacional, puesto que no existen pruebas reconocidas ni válidas en este ámbito (Mendoza, 2015); de hecho, los estudiantes extranjeros de intercambio universitario no cuentan con un certificado que acredite sus competencias en EFA. Acorde con esto, este estudio se plantea como objetivos: 1) Proponer un modelo de criterios evaluativos para el diseño de un prototipo de prueba de EFA, en línea, para un nivel B2 y 2) Aplicar los criterios evaluativos del modelo en la elaboración y validación de contenido del instrumento.

El artículo se organiza en dos apartados. En la sección 1 se abordan los fundamentos teóricos que sustentan este trabajo y en la sección 2, se describe el modelo de criterios y su operacionalización. Finalmente, se presenta las conclusiones y proyecciones de este trabajo.

1. Fundamentos teóricos

1.1. La evaluación del Español como Lengua Extranjera con fines académicos (EFA)

En la evaluación del EFA no está claro lo que una prueba de este tipo debería evaluar. Esta situación se debe a que no se han delimitado los contextos de uso que deberían abordarse, los cuales tienen que dar cuenta de las distintas habilidades que necesitan demostrar los candidatos o estudiantes en este ámbito. Es sabido que no se puede desarrollar un examen individualizado o dirigido a grupos restringidos, motivo por el cual se requiere de estudios que contribuyan a la determinación de los contextos de uso de lengua para fines académicos, puesto que las investigaciones actuales en pruebas de español con fines académicos no cumplen con los estándares definidos en los códigos de práctica por las distintas sociedades de evaluación (Llorián, 2012).

La evaluación del EFA, a diferencia del inglés, tiene un desarrollo muy incipiente. Esto explica que actualmente no haya ninguna prueba en este ámbito que tenga una validez nacional o internacional (Mendoza, 2015). Esto se debe, en parte, a que los estudiantes extranjeros de intercambio en universidades de habla hispana no deben acreditar mediante un certificado su dominio de español en contextos académicos, a diferencia del inglés o de otras lenguas extranjeras (Pastor, 2014). En este sentido, se requiere de esfuerzos desde las universidades y de los organismos gubernamentales para aunar criterios que ayuden a mejorar esta situación (Llorián, 2012).

1.2. Directrices para el diseño de pruebas de EFA

Los expertos en evaluación de segundas lenguas señalan que en la elaboración de una prueba de lengua se debe tener claridad respecto de las consideraciones éticas y de equidad, las especificaciones del instrumento, la selección de materiales, el diseño de tareas e ítems, el pilotaje y la validación (Alderson et al., 1998; Martínez, 2011; Figueras y Puig, 2013). Sin embargo, no siempre las instituciones aplican de manera rigurosa estos criterios.

En relación con los aspectos éticos, a partir de los años noventa ha habido una preocupación respecto de la atención que merecen estos

temas en la evaluación de lenguas extranjeras (Bordón, 2006). Esto ha llevado a una reflexión sobre las consecuencias y el impacto que tienen las pruebas de lenguas para los candidatos, para las instituciones y para la sociedad en general. De acuerdo con esto, las pruebas tienen que mostrar respeto y equidad por las personas y sustentarse en códigos de práctica que delimiten de manera objetiva los estándares necesarios para elaborar pruebas válidas.

Las especificaciones de un examen constituyen el plan en el cual se basan quienes redactan una prueba y sus ítems (Martínez, 2011). En ellas se establece: el propósito de la prueba, el perfil de los estudiantes o candidatos, los usuarios de la prueba, el o los niveles que evalúa, el marco de referencia teórico en el cual se sustenta, las habilidades que evalúa, los tipos de textos, los contenidos, los tipos de tareas e ítems, las ponderaciones, la duración, el formato de aplicación, entre otros (Alderson et al., 1998). Contar con la definición de cada una de estas especificaciones es relevante para el equipo durante todas las etapas de elaboración y validación de una prueba.

Respecto de la selección de materiales, es importante señalar que es complejo encontrar textos orales auténticos que se ajusten a los requerimientos de una prueba determinada. En este contexto, para los niveles superiores se sugiere conservar los ruidos de fondo, las reformulaciones, las repeticiones, etc. ya que se relacionan con el tipo de lengua que el candidato encuentra fuera del contexto del examen (Figuera y Puig, 2013).

En lo referido al diseño de tareas escritas y orales, se debe tener claridad sobre el tema que van a escribir o hablar los estudiantes, el cual debe estar acorde con el contexto de uso y el nivel de lengua definido en la prueba. En las tareas de escritura de EFA se debe evaluar el manejo de los distintos tipos de textos y sus formatos, la habilidad de escribir frases correctas desde una perspectiva gramatical, léxica, de puntuación y de ortografía; la habilidad para dar coherencia a las frases y párrafos y cohesión al texto (Martínez, 2011). La expresión e interacción oral, generalmente, se evalúa mediante la técnica de entrevista. Al respecto, señala Martínez (2011) que la entrevista debe ser lo más parecida a una conversación o actividad habitual que enfrenta el estudiante, para así evitar el nerviosismo del candidato. El tiempo mínimo para lograr muestras relevantes en una entrevista está fijado en 10 minutos. Esto

porque habrá estudiantes que al inicio serán más tímidos y si no se le da las oportunidades no se podrá observar su competencia real en la lengua (Bordón, 2006).

1.3. Ventajas de la implementación de pruebas de EFA a través del computador

En la actualidad, hay un gran interés por automatizar los exámenes de lenguas extranjeras o segundas lenguas. A partir de 1997 algunas instituciones internacionales de evaluación de lenguas extranjeras, como el English Testing Service (ETS), encargada de administrar el examen TOEFL, ha implementado sistemas automáticos en la evaluación del inglés como lengua extranjera. Es así como en una primera etapa se utilizó el formato Computer Based TOEFL y más recientemente el Internet Based TOEFL. No obstante, en el caso del español como lengua extranjera se han realizado pocos intentos de automatización de pruebas de proficiencia; de hecho, las pruebas de mayor reconocimiento internacional continúan aplicándose en formato de lápiz y papel. En este contexto, conviene preguntarse ¿Cuál es el aporte o la contribución de las pruebas automatizadas? Para García y Navarro (2008), la aplicación de pruebas en lápiz y papel tiene un costo elevado tanto para los estudiantes como para las instituciones. Este costo se observa en aspectos como el desplazamiento del candidato al centro donde se aplica la prueba, el proceso de entrega de los resultados, y la necesidad de contar con salas, examinadores, correctores y cuidadores. En este sentido, una prueba en línea representa una ventaja porque es menos costosa. Además, mediante este formato de implementación se tiene flexibilidad en cuanto a los diferentes tipos de ítems, se puede evaluar globalmente las destrezas comunicativas, dar una corrección rápida y eficaz, y las pruebas orales se pueden realizar sin la presencia de los evaluadores.

1.4. Validez de contenido en las pruebas de EFA

En el proceso de elaboración de una prueba de EFA, tan relevante como la automatización del instrumento es que éste sea válido estadísticamente para los fines que fue elaborado. En este sentido, la validación de contenido tiene como objetivo medir el grado en que una prueba evalúa lo que pretende evaluar (Alderson et al., 1998; Bordón, 2006; Martínez, 2011). Este tipo de validez consiste en solicitar a personas expertas en el

tema del estudio que realicen un juicio sobre un instrumento que se ha diseñado. El grupo de expertos evalúa si los ítems considerados en la elaboración de la prueba son indicadores de lo que se pretende evaluar con dicho instrumento (Alderson et al., 1998).

1.5. Estándares internacionales para la elaboración de pruebas

Las principales organizaciones en materia de evaluación de lenguas son: ALTE, ILTA y en el ámbito del ELE, el SICELE. Dichas entidades han tenido como objetivo común la elaboración de códigos de práctica que orientan respecto de las consideraciones éticas, de diseño y de impacto que hay que tener en cuenta al momento de elaborar y administrar pruebas de lengua. Los parámetros abordados por las distintas sociedades son indicaciones generales que luego cada institución o profesor aplican según sus necesidades y disponibilidad de recursos. Por lo tanto, se requiere de modelos que operacionalicen estos principios generales y provean criterios que ayuden a resolver de manera empírica la problemática sobre cómo elaborar y validar pruebas de EFA mediatizadas por computador.

1.6. Pruebas de proficiencia que certifican el dominio del EFA

En la actualidad, las pruebas de español general que tienen mayor reconocimiento a nivel internacional en la certificación del dominio son los DELE del Instituto Cervantes y el CELU del Consorcio Interuniversitario de Español como Lengua Segunda o Lengua Extranjera de Argentina. En el ámbito del EFA, se dispone de un número muy reducido de instrumentos y ninguno tiene un reconocimiento internacional. Se trata de los exámenes DUCLE (Diploma Universitario de Competencia en Lengua Española como lengua extranjera) de la Universidad Nacional de Rosario, del CELA (Certificado de Español como Lengua Adicional) y del EXELEAA (Examen de Español como Lengua Extranjera para el Ámbito Académico). Estos dos últimos elaborados por el Centro de Español para Extranjeros (CEPE) de la Universidad Nacional Autónoma de México. El DUCLE acredita el grado de dominio de español de los candidatos y puede ser presentado en instituciones educativas, gubernamentales o empresas privadas. Este examen certifica los niveles básico, intermedio y avanzado (estos dos últimos niveles están en el contexto de los ámbitos

laborales y académicos). El CELA evalúa el conocimiento y uso del español para interactuar en los ámbitos personal, académico y público, y certifica los niveles B1, B2 y C1 según el MCER (2002). El EXELEAA evalúa el nivel de dominio lingüístico de candidatos que necesitan acreditar competencia en la lengua para cursar estudios superiores.

De las pruebas mencionadas, solo es posible acceder a descripciones generales del instrumento y a algunas muestras de pruebas o reactivos; sin embargo, se desconoce cuáles fueron las consideraciones que se tuvieron en su diseño, elaboración y validación.

2. Metodología del estudio

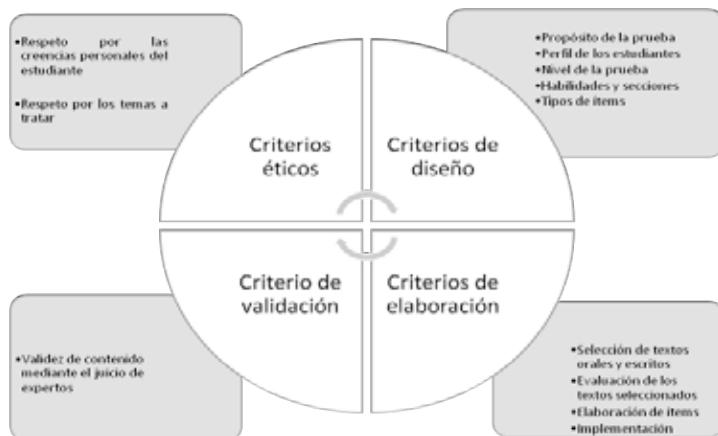
La investigación que se llevó a cabo, con la finalidad de desarrollar un prototipo de prueba de EFA, en línea, de nivel B2, corresponde a un estudio mixto de tipo exploratorio secuencial que consta de dos etapas:

- 1) **Cualitativa:** Consistió en la revisión de los criterios evaluativos propuestos por los expertos en elaboración de pruebas de segundas lenguas y de los códigos de práctica de las distintas sociedades evaluativas. Estos datos permitieron definir el modelo de criterios para diseñar un prototipo de prueba EFA en línea, como se observa en la figura 1.
- 2) **Cuantitativa:** Basado en el modelo de criterios se diseñó, desarrolló y validó a nivel de contenido el prototipo de prueba EFA en línea de nivel B2.

2.1. Etapa 1: Modelo de criterios evaluativos para el diseño del prototipo de prueba EFA

En relación con la problemática planteada, y acorde con el primer objetivo del estudio, se definió el modelo de criterios. Como se observa en la figura 1, este modelo se compone de criterios: 1) éticos, 2) de diseño, 3) de elaboración y 4) de validación. Se trata de los criterios necesarios para diseñar, elaborar y validar un prototipo de prueba EFA en línea de nivel B2 dirigida a estudiantes extranjeros de intercambio universitario. Es importante señalar que este modelo es interactivo, puesto que los criterios no se abordan solo de una manera lineal, sino que se retroalimentan entre sí. Por ejemplo, los criterios de validación pueden incidir en los criterios de diseño o de elaboración.

Figura 1: Modelo de criterios para el diseño, elaboración y validación del prototipo de prueba EFA



Los criterios éticos de este modelo son fundamentales, ya que se consideran desde el diseño hasta la validación del prototipo de prueba. Esto se refleja en la selección temática de los textos orales y escritos, y de las tareas abiertas que elicitán el dominio del español en contextos académicos; en todos ellos se muestra respeto por las creencias, la procedencia y la formación del estudiante.

Los criterios de diseño corresponden a las directrices que orientan la elaboración del instrumento. En el diseño se definió que el prototipo de prueba EFA se orienta a los estudiantes extranjeros de intercambio universitario y que, mediante ítems de corrección objetiva y de tareas abiertas, se evalúa las cuatro habilidades lingüísticas y la competencia lingüística en un nivel B2. En los criterios de elaboración se considera la selección de los materiales y su evaluación por un grupo de expertos, lo cual permite luego la redacción de los ítems y su implementación en línea. Finalmente, el criterio de validación contempla la participación de un grupo de expertos en el ámbito, quienes califican si los ítems y tareas están midiendo las dimensiones definidas.

De acuerdo con el tipo de estudio, en esta primera etapa se recopilaron y seleccionaron los datos cualitativos necesarios para elaborar el prototipo de prueba EFA en línea de nivel B2. En la etapa siguiente, se aplicaron los criterios del modelo en la elaboración y validación de contenido del instrumento.

2.2. Etapa 2: Aplicación del modelo de criterios evaluativos en la elaboración del prototipo de prueba EFA

En esta segunda etapa se presenta el desarrollo del modelo de criterios en la elaboración y validación del prototipo de prueba EFA en línea de nivel B2.

2.2.1. Criterios éticos

La operacionalización de los criterios éticos se consideró en la selección de los textos orales y escritos y en los temas de las tareas abiertas de expresión e interacción escrita y oral, donde se cauteló que las temáticas fueran propias de un ámbito académico. De esta manera, en el prototipo de prueba no se presenta temas polémicos, discriminatorios o que privilegien determinados contenidos disciplinares, pues lo que se evalúa está en el contexto de las habilidades lingüísticas y no del contenido temático.

2.2.2. Criterios de diseño

De acuerdo con los criterios de diseño de la figura 1, el prototipo de prueba EFA es una prueba de proficiencia que evalúa lo que es capaz de hacer (habilidades) el estudiante en cuanto al uso del español en contextos académicos y está orientada a un nivel B2¹. El perfil de los estudiantes extranjeros de intercambio, según los datos del Programa de Español como Lengua Extranjera de la Universidad de Concepción (ELE-UdeC), corresponde a alumnos de distintas áreas del conocimiento, siendo las más representativas: Letras (22%), Ingeniería (21%), Ciencias Jurídicas (16%), Economía (14%), Ciencias Físicas y Matemáticas (13%) y Ciencias Sociales (6%). Un estudiante de nivel B2 en un ámbito académico es capaz de comprender las ideas principales de textos complejos, de temas concretos y abstractos; interactuar con hablantes nativos con un grado suficiente de fluidez y naturalidad; producir textos claros y detallados sobre diversos temas; y defender un punto de vista sobre temas específicos, indicando los pros y los contras de las distintas opciones (MCER, 2002). Los contextos de uso de la lengua corresponden a los ámbitos propios

¹ El Marco Común Europeo de Referencia para las Lenguas, MCER (2002), define seis niveles en el dominio de una segunda lengua o lengua extranjera: A1 (Acceso), A2 (Plataforma), B1 (Umbral), B2 (Avanzado), C1 (Dominio operativo eficaz) y C2 (Maestría).

de un nivel B2 de EFA; esto es, participar de conversaciones sobre temas especializados; escuchar y comprender programas científicos, charlas, conferencias; leer y escribir sobre temas especializados.

En relación con las habilidades lingüísticas evaluadas, se elaboró una tabla de especificaciones donde se definió de manera operativa lo que cada una comprendía. A modo de ejemplo, se observa en la figura 2 que la expresión e interacción escrita se evalúa en función de la competencia discursiva, el dominio gramatical y del léxico del estudiante.

Figura 2: Tabla de especificaciones del prototipo de prueba EFA B2

Expresión e interacción escrita		
C. discursiva	Gramática	Vocabulario
Utiliza una variedad de palabras para enlazar frases y crear un texto claro y coherente, según el formato solicitado.	Tiene un buen dominio gramatical, aunque de manera escasa puede cometer errores no sistemáticos en la estructura de la frase, estos no producen malentendidos.	Tiene un amplio repertorio léxico que le permite expresarse sobre una variedad de temas. Si comete algún error de selección de léxico, este no afecta la comunicación.

Cada una de las habilidades se evalúa de manera separada, por lo que el prototipo de prueba se compone de cinco secciones, como se observa en la tabla 1.

Tabla 1: Secciones del prototipo de prueba EFA

Sección	Nº textos	Tipo y nº de ítems	Tiempo
Comprensión lectora (25%)	3 textos (fragmentos de artículos científicos)	V o F (1 texto con 5 preguntas) Selección múltiple con 4 opciones (2 textos con 5 preguntas cada uno)	30 minutos
Comprensión auditiva (20%)	3 audios (exposiciones)	V o F (1 audio con 5 preguntas) Selección múltiple con 4 opciones (2 audios con 5 preguntas cada uno)	30 minutos

Competencia lingüística (10%)	1 texto (ensayo)	Selección múltiple con 4 opciones	10 minutos
Expresión e interacción escrita (25%)	---	1 tarea abierta con dos opciones (el estudiante elige una)	20 minutos
Expresión e interacción oral (20%)	---	Entrevista: Consta de 3 partes: -Presentación -Elección del tema a exponer -Conversación sobre el tema	15 minutos
TOTAL			1 hora, 45 minutos

La tabla 1 muestra la distribución de las secciones del prototipo de prueba, sus porcentajes, el número y tipo de ítems y el tiempo asignado para cada una de las partes. Este prototipo de prueba en línea tiene una duración total de 1 hora, 45 minutos. Las secciones de comprensión lectora, auditiva, competencia lingüística y de expresión e interacción escrita se evalúan a través del computador, y la expresión e interacción oral de manera presencial una vez que el estudiante ha terminado la prueba en línea. Cabe señalar que esta habilidad también podría evaluarse a través de Skype si el estudiante se encontrara en su país de origen o no pudiera asistir a rendir la prueba de manera presencial.

Como se observa en la tabla 1, se contó con tres tipos de ítems distintos para evitar que el instrumento se orientara hacia un método o a un tipo de estudiante (Alderson et al., 1998). Los ítems de corrección objetiva para las secciones de comprensión lectora y auditiva fueron de verdadero y falso y de selección múltiple con cuatro opciones. La sección de comprensión lectora se compone de tres textos y la sección auditiva de tres audios. En el primer texto de cada una de estas secciones se plantean cinco preguntas de verdadero y falso, y en los textos 2 y 3 se plantean preguntas de selección múltiple (5 preguntas por cada texto). La sección de competencia lingüística comprende un texto con veinte ítems de selección múltiple donde se evalúan los diferentes usos gramaticales definidos para el nivel B2 en el contexto del EFA. Los ítems de respuesta abierta se consideraron para evaluar la expresión e interacción escrita y oral. En la sección de expresión e interacción escrita, el estudiante debe escribir un ensayo. Para ello, se propone dos temas distintos. El

alumno elige uno y escribe el texto según el formato entregado. En la expresión e interacción oral, el estudiante tiene que cumplir con tres tareas de la entrevista: 1) presentarse, 2) exponer sobre un tema a elección y argumentar un punto de vista sobre dicho tema y 3) conversar con el entrevistador sobre el tema expuesto.

2.2.3. Criterios de elaboración

Teniendo definido el diseño del prototipo de prueba EFA se procedió a la búsqueda y selección de los textos orales y escritos, acorde con el perfil de los estudiantes. En esta selección se consideró que los textos fueran éticamente adecuados; es decir, trataran de diversos temas acordes con un nivel B2 de EFA y con los datos del Programa ELE-UdeC, por lo que se recopilaron textos de diversas fuentes: revistas de divulgación académica, programas de radio y televisión, y entrevistas a expertos. Los textos corresponden a fragmentos de artículos científicos, charlas y ponencias, por lo que son representativas las modalidades expositivas y argumentativas. La longitud de los textos escritos seleccionados fue entre 300 y 500 palabras, y la de los orales, entre 1 y 3 minutos. En estos textos se utiliza un lenguaje y estilo académico, sin expresiones idiomáticas que dificulten la comprensión. En los audios se cauteló que hubiera una articulación y velocidad adecuadas para el nivel B2. En algunos casos se hicieron adaptaciones de los textos orales y escritos, manteniendo su organización discursiva (coherencia y cohesión). De acuerdo con los criterios de diseño, se recopiló un conjunto de textos, de los cuales finalmente se seleccionaron diez textos orales y diez escritos que se consideraron más adecuados en términos de pertinencia, contenido y formato para el nivel B2 de EFA. Luego, estos textos fueron evaluados por un grupo de diez expertos (5 profesores de segundas lenguas con grado de Magíster en Lingüística, 2 estudiantes de Magíster en Lingüística Aplicada y 3 Doctores en Lingüística) a través de una escala Likert con el objeto de validar en qué medida estos materiales eran aptos para evaluar el nivel de proficiencia B2 de EFA. La evaluación de los textos consideró tres criterios: pertinencia, contenido y formato, como se observa en la figura 3.

Figura 3: Criterios para evaluar los textos

I. Descripción de los criterios para evaluar los textos

CRITERIOS	INDICADORES
PERTINENCIA: El tipo de texto y el tema son adecuados para un nivel B2 de ELE	1. Muy en desacuerdo 2. En desacuerdo 3. De acuerdo 4. Muy de acuerdo
CONTENIDO: El contenido gramatical y léxico es adecuado para un nivel B2 de ELE	1. Muy en desacuerdo 2. En desacuerdo 3. De acuerdo 4. Muy de acuerdo
FORMATO DEL TEXTO: La redacción, el tamaño de letra, la longitud del texto/ la locución, el ruido de los audios son apropiados para el nivel que se orienta la prueba	1. Muy en desacuerdo 2. En desacuerdo 3. De acuerdo 4. Muy de acuerdo

Acorde con el juicio de expertos, se seleccionaron los tres textos que presentaban un 80% de acuerdo entre los evaluadores en los tres criterios definidos (tres textos para la comprensión lectora y tres para la comprensión auditiva). Los textos seleccionados para la comprensión lectora corresponden a tres fragmentos de artículos científicos: 1) El sistema de competencias sustentables para el desempeño profesional en ingeniería, 2) El aprendizaje de una segunda lengua y 3) La biodiversidad. Los textos para la comprensión auditiva corresponden a dos exposiciones de trabajos científicos-tecnológicos y un fragmento de un programa televisivo sobre ciencia: 1) La energía eólica, 2) Un proyecto para evitar la escasez de agua en la zona norte de Chile y 3) Los biocombustibles.

Los ítems de las secciones de comprensión lectora y auditiva se elaboraron teniendo en cuenta tres dimensiones o microhabilidades: 1) la identificación de información específica en textos especializados, 2) la inferencia de las ideas principales de los párrafos y 3) la inferencia de la idea global del texto (MCER, 2002). Los ítems de la sección de competencia lingüística miden la dimensión del dominio gramatical correspondiente a un nivel B2 de EFA. Las tareas de expresión escrita y oral evalúan dos dimensiones: 1) las modalidades expositiva y argumentativa y 2) el género académico. En consecuencia, el prototipo de prueba evalúa seis dimensiones en el dominio del EFA para el nivel B2.

En el proceso de elaboración de los ítems de comprensión lectora y auditiva se registró, junto a la pregunta planteada, la dimensión que cada una evaluaba, como se observa en la figura 4. Esto, sin duda, facilitó posteriormente la validación del prototipo de prueba por parte del grupo de expertos.

Figura 4: Ejemplo de elaboración de preguntas

PREGUNTA 5 Dimensión: Inferencia de la idea global del texto	
5. ¿Cuál es la idea central del texto?	
a) Los factores que intervienen en el aprendizaje de un segundo idioma	
b) Los diferentes factores que se adaptan al aprendizaje de una segunda lengua	
c) Las aptitudes de los estudiantes que aprenden un segundo idioma	
d) Los componentes necesarios para el aprendizaje efectivo de una segunda lengua	

En relación con la elaboración de tareas abiertas, se consideraron dos temas para la escritura y dos para la expresión oral. En la expresión e interacción escrita se consideraron la formación académica a distancia y El día Mundial sin compras. De estos, el estudiante selecciona uno y realiza la tarea en el tiempo asignado, como se observa en la figura 5.

Figura 5: Tarea de escritura

Sección IV: Expresión e interacción escrita
La sección de expresión escrita consta de dos opciones. Solo tienes que elegir UNA opción y escribir un ensayo de 250 palabras (20-25 líneas)
Opción 1 Una alternativa a la formación académica tradicional es tomar cursos a distancia (por Internet), modalidad que permite un acceso real al conocimiento de personas que antes no podían hacerlo, ya sea por razones económicas o de tiempo. La modalidad de aprendizaje a distancia, cada vez va tomando más fuerza, pero aún queda el escepticismo de algunos que dudan si es efectiva o no. En relación con este tema, escribe un ensayo sobre la formación académica a través de cursos a distancia por Internet. En el ensayo tienes que incluir: <ul style="list-style-type: none">• Escribir un título• Introducir el tema• Establecer una postura respecto del tema• Mencionar las ventajas y desventajas de la formación a través de internet• Concluir

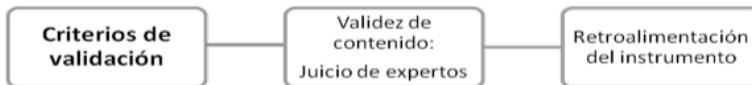
La evaluación de la expresión e interacción oral se lleva a cabo mediante la técnica de entrevista, compuesta de tres partes: 1) presentación del estudiante, 2) exposición y argumentación sobre un tema a elección (opción 1: el uso de pesticidas en la agricultura, opción 2: La inmersión o el aprendizaje formal de un segundo idioma) y 3) conversación con el entrevistador sobre el tema expuesto. Esta sección tiene una duración de 15 minutos.

El prototipo de prueba EFA de nivel B2 se implementó en la plataforma Moodle (del inglés *Modular Object-Oriented Dynamic Learning Environment*) con cuatro secciones. La sección cinco de expresión e interacción oral se realiza de manera presencial. Las opciones avanzadas de Moodle permiten el registro de los estudiantes y la programación de la prueba en cuanto a la fecha, la hora y la duración. El formato de aplicación de este prototipo de prueba representa un aporte a la evaluación del EFA, dado que tanto las pruebas generales de ELE como las pruebas de EFA, normalmente, se aplican en formato de lápiz y papel. En este sentido resulta beneficioso obtener de manera rápida y eficaz los resultados sobre el nivel de proficiencia de los estudiantes. Además, Moodle otorga seguridad en diversos aspectos: restringe el número de intentos del estudiante, por lo que una vez que ha comenzado la prueba no puede detenerla; muestra el tiempo del que dispone el alumno; facilita la navegación, ya que para avanzar a la sección siguiente solo se presiona el botón “siguiente”; en la misma pantalla el alumno dispone del texto y de las preguntas (a diferencia del papel donde tiene el texto y la hoja de respuestas). Del mismo modo, en la sección de competencia lingüística el alumno va leyendo el texto y va seleccionando la forma gramatical adecuada sin tener que salirse del texto para marcar la alternativa correcta. En cuanto a la escritura, la plataforma da la opción de desactivar el procesador de texto, lo cual permite acceder a muestras de lengua auténticas. Sin duda, este formato promueve una interacción entre el estudiante y la prueba (Figueras y Puig, 2013) y permite dos formas de aplicación: presencial y a distancia.

2.2.4. Criterios de validación

Este modelo consideró la validez de contenido del prototipo de prueba EFA a través del juicio de expertos, como se observa en la figura 6.

Figura 6: Criterios de pilotaje



En este proceso de evaluación participaron ocho expertos, profesionales de lenguas (2 estudiantes de Magíster en Lingüística Aplicada, 1 Magíster en Lingüística, 3 candidatos al grado de Doctor en Lingüística y 2 Doctores en Lingüística). De manera individual, cada experto evaluó, a través de una escala Likert de 5 puntos (1: Muy en desacuerdo, 2: En desacuerdo, 3: Indiferente, 4: De acuerdo, 5: Muy de acuerdo), en qué medida los ítems de las diferentes secciones del prototipo de prueba, implementado en Moodle, medían las dimensiones definidas en los criterios de elaboración. Para ello, se les entregó una pauta (figura 7) donde se le explicaba en qué consistía la evaluación y luego se proporcionaba la lista de ítems en las distintas dimensiones. Los evaluadores leían la pauta y luego accedían al prototipo de prueba en línea, con una clave que se les entregó. Enseguida leían los textos y evaluaban si los ítems medían las distintas dimensiones, y así sucesivamente con cada sección de este prototipo de prueba.

En la figura 7 se observa parte del instrumento que respondieron los evaluadores, específicamente las preguntas 3 y 4 correspondientes al primer texto de comprensión lectora (Sistemas de competencias sustentables para el desempeño profesional en Ingeniería) que evalúan la dimensión II del prototipo de prueba.

Figura 7: Pauta de evaluación de contenido para los jueces

DIMENSIÓN II: INFIERE LAS IDEAS PRINCIPALES DE LOS PÁRRAFOS

Secciones: Comprensión de lectura	Evalúa Dimensión	Indicadores				
Preguntas	SI/NO (observación)	1	2	3	4	5
<i>Texto I</i>						
3. Las competencias de egreso se reflejan en el dominio disciplinar que debe mostrar todo profesional de ingeniería.						
4. Se puede inferir que las escuelas de Ingeniería han intentado implementar cambios en la formación académica de sus estudiantes.						

Una vez que se completó la evaluación de los expertos, se tabularon los datos en Excel atendiendo a las calificaciones de los evaluadores. Los resultados obtenidos muestran que los ítems del prototipo de prueba presentan validez de contenido, en tanto que miden las dimensiones definidas. En la dimensión I de la comprensión lectora y auditiva, que corresponde a las preguntas 1 y 2 de cada texto, todos los jueces calificaron los ítems entre 4 y 5 en la escala Likert. En la dimensión II de las mismas secciones, que corresponden a las preguntas 3 y 4 de comprensión lectora y auditiva, los ítems también fueron calificados entre 4 y 5 en la escala Likert, excepto la pregunta 4 del texto 2, que dos expertos la calificaron con 3, dado que la forma en que estaba redactada no evaluaba la inferencia a nivel de párrafo. En la figura 8 se observan los resultados de la dimensión III, correspondiente a las preguntas 5 de cada texto escrito y oral. Estas preguntas evalúan la inferencia de la idea central del texto/audio.

Figura 8: Resultados juicio de expertos

DIMENSIÓN III: INFERENCIA DE LA IDEA CENTRAL DEL TEXTO					
Secciones:	Indicadores				
Comprensión lectora y auditiva	1	2	3	4	5
<i>Texto I</i>					
5. A partir de la lectura del texto se puede inferir que la se				1	7
<i>Texto II</i>					
5. ¿Cuál es la idea central del texto?				3	5
<i>Texto III</i>					
5. ¿Cuál es la idea central del texto?				3	5
<i>Audio I</i>					
5. La idea central del audio es dar cuenta del aporte de las					8
<i>Audio II</i>					
5. ¿Cuál es la idea central del audio?				1	7
<i>Audio III</i>					
5. ¿Cuál es la tema principal del audio?				2	6

Los resultados del juicio de expertos en la dimensión III del prototipo de prueba evidencian que las preguntas planteadas evalúan la dimensión que dicen evaluar. La dimensión IV del instrumento, que corresponde al dominio de la competencia lingüística para el nivel B2 de EFA, presenta un alto nivel de acuerdo entre los evaluadores (6 muy de acuerdo y 2 de acuerdo). En la dimensión V, correspondiente a si las tareas abiertas de expresión e interacción escrita y oral evalúan las modalidades expositivas y argumentativas, los evaluadores las calificaron como muy de acuerdo. La dimensión VI que evalúa si las tareas de expresión e interacción escrita y oral miden el género académico ensayo y exposición oral respectivamente, fue calificada como muy de acuerdo, y solo hicieron sugerencias en cuanto a abordar otros temas más académicos en la opción 2 de la tarea de escritura (El Día Mundial sin compras). Finalmente, se hicieron las modificaciones sugeridas por los expertos con la finalidad de mejorar la calidad y validez del prototipo de prueba. De esta forma, la opción 2 de escritura ahora se refiere a la crisis del agua a nivel mundial.

Los resultados obtenidos a través de este juicio de expertos permiten sostener la efectividad del modelo de criterios, dado que para la elaboración del prototipo de prueba y sus ítems fue esencial contar con

los criterios definidos en el diseño, lo cual permitió organizar de manera secuencial y rigurosa el proceso de desarrollo del instrumento. Este juicio de expertos ha servido para corroborar que los ítems y tareas son apropiados para medir el nivel de proficiencia B2 en EFA en un contexto mediatizado por computador.

3. Conclusiones

Teniendo en consideración que no hay pruebas de EFA reconocidas a nivel nacional o internacional, y que existe la necesidad de contar con un instrumento que permita evaluar a los estudiantes extranjeros en cuanto a sus competencias académicas, este estudio ha permitido responder a los dos objetivos planteados, dado que se ha propuesto el modelo de criterios para el diseño de un prototipo de prueba EFA en línea para el nivel B2 y estos criterios se han aplicado a la elaboración, desarrollo y validación de contenido del instrumento.

La relevancia de este estudio radica en que el prototipo de prueba EFA relaciona lo teórico y lo empírico para dar respuesta a cómo evaluar el español como lengua extranjera con fines académicos. Es necesario señalar que el formato de aplicación del instrumento y los temas abordados en los textos y tareas están acordes con el perfil de los estudiantes extranjeros definidos en el diseño. El entorno de la prueba es interactivo, de fácil navegación, donde el estudiante puede ser evaluado cuando aún se encuentra en su país de origen y el profesor puede acceder de forma rápida a los resultados. De esta manera, se pueden planificar estrategias que apoyen a los estudiantes en su desempeño académico en lo que respecta al uso del español.

El modelo propuesto resulta interesante, puesto que evidencia desde los procesos iniciales de planeación del prototipo de prueba en línea hasta las etapas finales de elaboración y validación, cuestión que los expertos reclaman a las instituciones que actualmente ofrecen exámenes de EFA (Llorián, 2012). No obstante, hay que señalar que la validación de una prueba no es algo que finaliza cuando ésta se ha administrado y evaluado, sino que se trata de un proceso que se realiza continuamente (Bordón y Liskin-Gasparro, 2007), y es lo que se hace en pruebas de alto impacto a nivel internacional como el TOEFL o los exámenes de Cambridge, donde ha habido mucho trabajo e inversión de parte de las instituciones patrocinantes para contar con instrumentos válidos y

confiables. Y es hacia allá donde se necesita avanzar en las pruebas de proficiencia de EFA.

Una de las proyecciones de este trabajo es realizar la validación de apariencia y el pilotaje con una muestra de estudiantes extranjeros universitarios que necesitan integrarse con fines académicos a nuestro país. Esto, con la finalidad de tener una mirada complementaria a la evaluación de los expertos. También sería interesante realizar una evaluación en cuanto a la usabilidad y funcionamiento del entorno computacional en que se encuentra implementada. Esta es una tarea compleja que exige la participación de equipos interdisciplinarios, porque considera tanto lo evaluativo como lo informático (García y Navarro, 2008).

Como limitación de este trabajo, podemos mencionar que es difícil proporcionar un modelo de elaboración y validación de pruebas que considere todas las variables y pueda ser replicado sin modificaciones. Puesto que son múltiples los escenarios y realidades educativas, el modelo puede ser mejorado o adaptado en función de las necesidades del usuario. En este estudio, el modelo de criterios y su aplicación ha permitido elaborar un prototipo de prueba EFA en línea de nivel B2 con validez de contenido para evaluar niveles de proficiencia en estudiantes extranjeros de intercambio. No obstante, esta propuesta puede ser adaptada acorde con las diferentes necesidades y realidades evaluativas de EFA.

Bibliografía

- Alderson, Ch., Clapham, C. y Wall, D. (1998). *Exámenes de idiomas*. Madrid: Cambridge University Press.
- ALTE. (2009). *Código de Práctica*. Recuperado de: http://www.alte.org/attachments/files/code_practice_es.pdf
- Bordón, T. (2006). *La evaluación de la lengua en el marco de E/L2: Bases y procedimientos*. Madrid. Editorial: Arcolibros.
- Bordón, T y Liskin-Gasparro, J. (2007). Evaluación. En M. Lacorte (Ed.). *Lingüística aplicada del español* (pp. 211-252). Madrid: Arco/Libros..
- Consejo de Europa. (2002). *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Madrid: Anaya.
- Ferreira, A., Vine, A y Elejalde, J. (2013). Hacia una prueba de nivel en Español como Lengua Extranjera. *Revista de Lingüística Teórica y Aplicada* 51(2), 73-103.
- Figueras, N. y Puig, F. (2013). *Pautas para la evaluación del español como lengua extranjera*. Madrid: Edinumen.
- García, J. y Navarro, C. (2008). Automatizar los exámenes del Diploma de Español como Lengua Extranjera (DELE): ¿Utopía o realidad? *Revista de Lingüística Teórica y Aplicada* 46(2), 81-93.
- ILTA. (2000). *Código de Ética*. Recuperado de: <http://ildaonline.com/index.php/enUS/resources/ilda-code-of-ethics>
- Llorián, S. (2012). Los retos de la certificación del español con fines específicos. *Revista de Nebrija de Lingüística Aplicada*, 12(6), 3-34.
- Martínez, A. (2011). *La evaluación de las lenguas: garantías y limitaciones*. Granada: Ediciones Mágina.
- Mendoza, A. (2015). La selección de tareas de escritura en los exámenes de lengua extranjera destinados al ámbito académico. *Revista de Nebrija de Lingüística Aplicada*, 18(9) ,1-18.

Pastor, S. (2014). Hacia una certificación oficial del español académico.

II Congreso Internacional del SICELE. Avances en la certificación del ELE, México D.F.

SICELE. (2015). *Estándares de calidad SICELE*. Recuperado de: [https://sites.google.com/a/sicelle.org/sicelle/politica-de-](https://sites.google.com/a/sicelle.org/sicelle/politica-de-calidad/estandares-sicelle)

[calidad/estandares-sicelle](https://sites.google.com/a/sicelle.org/sicelle/politica-de-calidad/estandares-sicelle)