

## Selección y diseño de tareas e ítems en las pruebas del SIELE

### *Selection and design of tasks and items in SIELE tests*

INÉS SORIA / AURORA DUQUE / JAVIER FRUNS / FRANCISCO LÓPEZ

Unidad de Certificación Lingüística. Departamento de Acreditación y Certificación. Dirección Académica. Instituto Cervantes

isoria@cervantes.es / aurora.duque@cervantes.es / jfrunsg@cervantes.es  
/ paco.lopez@cervantes.es

**Resumen:** La Unidad de Certificación Lingüística del Instituto Cervantes ha participado en el diseño del SIELE, Servicio Internacional de Evaluación de la Lengua Española. Este artículo:

- analiza las características de este nuevo sistema certificativo en contraste con los DELE;
- explica los criterios que se han utilizado en el diseño de las pruebas de este examen por ordenador;
- y describe el proceso de creación de tareas, en el que participan las instituciones promotoras, Universidad Nacional Autónoma de México, Universidad de Salamanca, Universidad de Buenos Aires e Instituto Cervantes, junto con Telefónica Educación Digital.

**Palabras clave:** SIELE, DELE, evaluación certificativa, evaluación asistida por ordenador

**Abstract:** The Language Certification Unit of the Instituto Cervantes has participated in the design of SIELE, International Service of Evaluation of the Spanish Language. This article analyzes the characteristics of this new certification system in contrast to the DELE system. It explains the criteria that have been used in the design of this computer test and describes the process of task creation, involving the promoting institutions (Universidad Nacional Autónoma de México, Universidad de Salamanca, Universidad de Buenos Aires e Instituto Cervantes) with Telefónica Educación Digital.

**Key words:** SIELE, DELE, Certification assessment, computer-aided assessment



El SIELE, Servicio Internacional de Evaluación de la Lengua Española, es un sistema de certificación de dominio del español a través de medios electrónicos, promovido por el Instituto Cervantes, la Universidad Nacional Autónoma de México (UNAM), la Universidad de Salamanca (USAL) y la Universidad de Buenos Aires (UBA) y gestionado por Telefónica Educación Digital.

En este artículo nos proponemos contextualizar el SIELE dentro de la oferta certificativa del Instituto Cervantes y resumir los trabajos en los que ha participado la Unidad de Certificación Lingüística de la Dirección Académica del Instituto Cervantes, como integrante de los equipos técnico-académicos del SIELE.

Para ello, en primer lugar contrastaremos el SIELE con el sistema certificativo de los Diplomas de español DELE, títulos oficiales de español del Ministerio de Educación Cultura y Deporte, cuya gestión administrativa, económica y académica corresponde al Instituto Cervantes.

A continuación, veremos los criterios que se han empleado en el diseño de las tareas y los procedimientos que se han aplicado para validar la diferentes pruebas que constituyen el SIELE.

Por último, describiremos el proceso de producción de las pruebas y tareas del SIELE.

## **1. DELE Y SIELE**

En el año 2009, la Dirección Académica del Instituto Cervantes inició el proceso de vinculación del sistema certificativo de los DELE con los niveles (A1, A2, B1, B2, C1 y C2) establecidos en el *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación* (MCER), tal y como se indicaba en el Real Decreto 264/2008, del 22 de febrero.

Dicha encomienda la asumió la Unidad de Certificación Lingüística, ya que es la unidad de la Dirección Académica que se ocupa del diseño de exámenes, de la producción de tareas y de la formación de los calificadores de la prueba de Expresión e interacción orales de todos los DELE y de los calificadores de la prueba de Expresión e interacción escritas de los dos exámenes destinados al público escolar.

Durante cinco años, se trabajó muy intensamente para, en primer lugar, crear tres nuevas certificaciones, el DELE A1, el DELE A2 y el DELE C1, pues veníamos de un sistema certificativo articulado en tres niveles de dominio: Inicial (B1), Intermedio (B2) y Superior (C2). Después se actualizaron los exámenes existentes, el DELE B1, el DELE B2 y el DELE C2 para vincularlos al MCER siguiendo los procedimientos establecidos en el *Manual for relating Language Examinations to the Common European Framework of Reference for Languages (CEFR)*, publicado por el Consejo de Europa en enero de 2009.

Una vez terminado el proceso de renovación del sistema de certificación DELE, ya articulado en los seis niveles de MCER, se crearon entre 2014 y 2015 dos nuevos exámenes, el DELE A1 para escolares y el

DELE A2/B1 para escolares, diseñados para niños y jóvenes entre 11 y 17 años.

Por lo tanto, el sistema certificativo de los DELE cuenta en estos momentos con seis diplomas DELE, uno por cada nivel del MCER, que se obtienen a través de un sistema de ocho exámenes: seis de carácter general y dos dedicados al público escolar.<sup>1</sup>

Los exámenes DELE son exámenes de dominio de un nivel del MCER, excepto el A2/B1 para escolares que es un examen con doble salida: dependiendo del desempeño en este examen, el candidato puede obtener las calificaciones de No apto, Apto A2 o Apto B1. Así pues, los exámenes DELE no nos dicen el nivel que tiene el candidato sino que evalúan si el candidato alcanza el nivel de dominio que quiere certificar.

Todos los exámenes DELE constan de cuatro pruebas (excepto el DELE C2 que tiene tres): Comprensión de lectura, Comprensión auditiva, Expresión e interacción escritas y Expresión e interacción orales. A partir del examen B2, se inicia un proceso de integración de destrezas en las pruebas, de forma que podemos encontrar, por ejemplo, una grabación en audio como estímulo para la realización de una de las tareas de escritura. En el examen DELE C2, esta integración es prácticamente completa, como podemos apreciar en el nombre de las pruebas que lo componen:

- Prueba 1. Uso de la lengua, comprensión de lectura y auditiva
- Prueba 2. Destrezas integradas: comprensión auditiva y expresión e interacción escritas
- Prueba 3. Destrezas integradas: comprensión de lectura y expresión e interacción orales

En relación con las variedades del español que podemos encontrar en los exámenes, también hay una gradación dependiendo de los niveles. En los exámenes DELE A1 y A2, se emplean de variedades del español peninsular contemporáneo, tanto en los textos de entrada escritos como orales. A partir del DELE B1, se emplean textos de diferentes variedades del español, no solo peninsular. En los textos que produce el candidato, será considerada válida toda norma lingüística hispánica seguida coherentemente y respaldada por grupos amplios de hablantes cultos.

Vemos, por lo tanto, que el sistema certificativo de los DELE abarca el dominio del español en todos los niveles establecidos en el MCER en las diferentes actividades de la lengua. Llegados a este punto, podríamos preguntarnos por qué le podía interesar al Instituto Cervantes desarrollar otro tipo de certificación.

Desde el año 2011, el Instituto Cervantes buscaba la forma de impulsar un gran proyecto de colaboración con otras instituciones del mundo hispánico. De esta forma se empezó a colaborar con la Universidad

---

<sup>1</sup> En <http://exámenes.cervantes.es> se encuentra toda la información disponible sobre los DELE.

Nacional Autónoma de México y la Universidad de Salamanca y se decidió iniciar esa andadura con un nuevo sistema de evaluación de la lengua española.

Este nuevo sistema no es, por lo tanto, únicamente del Instituto Cervantes, sino que ha sido creado y desarrollado en colaboración con la UNAM, la USAL y la UBA, que se incorporaría al proyecto más adelante.

Escuchamos frecuentemente que los exámenes son herramientas que nos permiten medir el nivel de competencia de los candidatos pero, en realidad, son pruebas que nos permiten realizar hipótesis sobre las posibles actuaciones lingüísticas de los candidatos en diferentes contextos de la vida real. En este sentido, podríamos compararlas con distintas las pruebas médicas que permiten realizar un diagnóstico sobre nuestra salud: con pruebas diferentes podemos llegar a diagnósticos semejantes, si bien cada prueba es apropiada para momentos, necesidades y finalidades distintas.

Del mismo modo, las pruebas del DELE están diseñadas para obtener la mayor cantidad de información posible sobre la competencia del candidato en relación con un determinado nivel: busca todos los indicios, el mayor número de evidencias posible. El diploma DELE de un nivel garantiza que el candidato domina ese nivel del MCER en todas sus dimensiones.

En el SIELE, sin embargo, se buscan los elementos clave que permitan elaborar una hipótesis sobre la competencia lingüística del candidato en relación con los niveles del MCER.

Los cuatro pilares fundamentales sobre los que se acordó llevar a cabo el diseño de este nuevo sistema son los siguientes:

- constaría de un único examen, desarrollado de manera conjunta por equipos e instituciones de diferentes países del mundo hispanohablante;
- el examen sería multinivel y permitiría ubicar la actuación del candidato en un nivel de dominio del MCER;
- el examen englobaría las diferentes variedades del español.
- la realización del examen y la gestión académica del mismo se deberían realizar por ordenador exclusivamente.

Este punto de partida suponía para la Unidad de Certificación Lingüística un conjunto de retos apasionantes desde el punto de vista de la evaluación certificativa: debíamos trasladar nuestra experiencia en el campo de la evaluación lingüística a un entorno y soporte completamente diferente, el de la evaluación en línea mediante ordenador, y debíamos hacerlo sin renunciar a la evaluación de la expresión oral.

Otras decisiones que se tomaron al inicio del proyecto y que condicionaban el diseño de las pruebas fueron las siguientes:

- Se acordó que se desarrollarían cuatro pruebas, correspondientes a las actividades de la lengua de comprensión

de lectura, comprensión auditiva, expresión e interacción escritas y expresión e interacción orales.

- Se consideró la necesidad de que se pudieran realizar diferentes modalidades, en las que el candidato se presentara a algunas pruebas del examen en combinaciones predefinidas, para crear certificaciones parciales de competencia en las diferentes actividades de la lengua.
- Se estableció que, en una primera fase, el examen sería multinivel progresivo, no adaptativo, lo que implicaba que en todas las pruebas debía haber ítems y tareas de todos los niveles.
- Se determinó que la duración máxima del examen completo sería de tres horas, por razones de viabilidad y de atención a la capacidad de concentración de los candidatos.

Desde la perspectiva del Instituto Cervantes, un examen multinivel permitía complementar el esquema de certificación de dominio por niveles de los DELE. Además, se nos brindaba la oportunidad de dar respuesta a ciertas demandas que el DELE no podría cubrir, relacionadas principalmente con la rapidez en la obtención de resultados y con la posibilidad de certificar competencias parciales en relación con las actividades de la lengua.

## **2. Primera fase de diseño de las pruebas y tareas del SIELE**

En un primer momento, los equipos técnico-académicos definieron el constructo de cada prueba, tomando la determinación de utilizar el MCER y el *Plan Curricular del Instituto Cervantes. Niveles de referencia* como punto de partida. Participaron 15 expertos en evaluación de las instituciones promotoras y se trabajó por videoconferencia durante unos tres meses para desarrollar el cuadro de especificaciones generales.

Se decidió seleccionar tareas prototípicas de cada uno de los niveles del MCER, del A1 al C1, que son los niveles que cubre el examen. Todas las instituciones participantes han desarrollado exámenes certificativos, por lo que se llevó a cabo un proceso de selección de aquéllas tareas que ofrecían mejores resultados psicométricos. Seguramente si se leen las especificaciones de las pruebas se podrán identificar algunas tareas que forman parte del DELE, por ejemplo. Buscamos, en definitiva garantizar la validez y fiabilidad de nuestra prueba basándonos en la experiencia de todos los participantes.

Una vez seleccionadas las tareas, se realizaron los ajustes necesarios para que cada una de las pruebas resultase equilibrada y la cobertura del examen fuera lo más amplia posible, dentro de las limitaciones de tiempo establecidas. Se ajustaron las tareas, se seleccionaron unas y se desecharon otras para tratar de combinar:

- en las dos pruebas de comprensión y en la prueba de expresión e interacción orales, al menos una tarea de cada uno de los niveles que se contemplaban;
- la evaluación de diferentes capacidades de comprensión oral y auditiva en cada una de las tareas que componen estas dos pruebas, como por ejemplo captar la idea global, localizar información específica, identificar detalles, etc.;
- diferentes situaciones de comunicación y tipos de texto, siguiendo los descriptores del MCER y los inventarios del PCIC: anuncios, correos electrónicos, conferencias, etc.;
- diferentes ámbitos, aunque teniendo en cuenta que el SIELE nace con una clara vocación académica, por lo que en las tareas de B2 y C1 se encuentran actividades propias de ese contexto.

Como resultado de este proceso de selección y depuración, se confeccionó un prototipo, que llamamos Modelo 0, en el que quedaba plasmada la estructura de las pruebas tal y como son actualmente. Ese prototipo se experimentó en papel con más de 100 candidatos y, tras analizar los resultados, se ajustaron los tiempos, las instrucciones y la composición final de examen. Reproducimos a continuación el Cuadro de especificaciones del SIELE tal y como se presenta en el Anexo I de la *Guía oficial para preparadores del SIELE* (en prensa).

PRUEBA 1. COMPRENSIÓN DE LECTURA			
60 min 38 ítems	FOCALIZACIÓN Capacidad que se evalúa	FORMATO Tarea que se realiza	TEXTOS Características de los textos
<b>Tarea 1 5 ítems (A1)</b>	Extraer ideas generales e identificar información concisa y explícita en textos breves y sencillos.	Leer cinco textos y responder a cinco preguntas (dos referidas a ideas generales y tres sobre informaciones más explícitas) con tres opciones de respuesta cada una.	Mensajes, correos electrónicos o textos adaptados de material informativo o promocional (artículos muy breves y sencillos de revistas o periódicos, folletos, anuncios publicitarios, convocatorias, blogs, foros y avisos de acontecimientos de diverso tipo). Ámbito personal y público. Extensión: 40-80 palabras cada texto.
<b>Tarea 2 5 ítems (A2)</b>	Extraer las ideas principales e identificar información específica, sencilla en un texto extenso.	Leer un texto y responder a cinco preguntas (dos referidas a la idea principal y tres sobre informaciones más explícitas) con tres opciones de respuesta cada una.	Cartas y correos electrónicos personales y otros textos epistolares breves de uso habitual, como solicitudes de información sencilla, reservas, peticiones de aclaración sobre plazos, horarios o modalidades de pago. Ámbito personal y público. Extensión: 250-300 palabras.
<b>Tarea 3 8 ítems (B1)</b>	Localizar información específica en textos descriptivos, narrativos o informativos.	Leer tres textos y relacionarlos con ocho enunciados o preguntas.	Anécdotas, información práctica de guías de viajes, experiencias, noticias, diarios, biografías, ofertas de trabajo. Ámbito público. Extensión: 100-120 palabras cada texto.
<b>Tarea 4 8 ítems (B2)</b>	Reconstruir la estructura de textos e identificar las relaciones entre las ideas incluidas en ellos.	Leer dos textos incompletos, con cuatro huecos cada uno, e identificar, entre cinco fragmentos posibles para cada texto, los que corresponden a los espacios.	Artículos de opinión, noticias, cartas al director, guías de viaje... Ámbitos público, educativo y profesional. Extensión: dos textos de 230-280 palabras cada uno.
<b>Tarea 5 12 ítems (2 ítems de B2 y 10 ítems de C1)</b>	Identificar las estructuras, léxico y mecanismos de cohesión adecuados en textos con un repertorio lingüístico complejo.	Leer un texto con 12 espacios y elegir una de las tres opciones disponibles para cada espacio. Distribución: cuatro ítems centrados en identificar estructuras gramaticales (uno de ellos de nivel B2), cuatro en el léxico (uno de ellos de nivel B2) y cuatro en los mecanismos de cohesión.	Textos largos y complejos que tratan de temas especializados, extraídos de revistas, libros de texto y periódicos especializados. Ámbitos educativo y profesional. Extensión: 375-425 palabras.

PRUEBA 2. COMPRENSIÓN AUDITIVA			
55 min 38 ítems	FOCALIZACIÓN Capacidad que se evalúa	FORMATO Tarea que se realiza	TEXTOS Características de los textos
Tarea 1 5 ítems (A1)	Captar aspectos concretos en conversaciones a velocidad lenta y con articulación clara.	Escuchar dos veces una conversación y completar cinco enunciados breves con información del texto de entrada disponible en un banco común con quince opciones.	Conversación entre dos interlocutores que tratan sobre temas conocidos o cotidianos. Ámbito personal y público. Extensión: 150-170 palabras (máx. 2 minutos).
Tarea 2 5 ítems (A2)	Captar la idea general e información específica en noticias radiofónicas o fragmentos breves y sencillos.	Escuchar cinco monólogos breves dos veces y responder a cinco preguntas con tres opciones de respuesta cada una.	Cinco noticias o cortes radiofónicos de extensión media y estructura sencilla que tratan aspectos cotidianos y previsibles. Ámbito público. Extensión: 250-300 palabras en total (máx. 3 minutos).
Tarea 3 8 ítems (B1)	Captar la idea general de monólogos o conversaciones breves informales.	Escuchar ocho monólogos dos veces y relacionarlos con los enunciados que les corresponden, de un total de once disponibles.	Ocho monólogos o conversaciones de carácter informal en los que se cuentan anécdotas o experiencias personales sobre un mismo tema. Ámbitos público y profesional. Extensión: 50-70 palabras cada audición (máx. 5 minutos).
Tarea 4 8 ítems (B2)	Extraer información concreta y detallada e inferir posibles implicaciones en una conversación.	Escuchar una entrevista dos veces y contestar a ocho preguntas con tres opciones de respuesta cada una.	Entrevista radiofónica o televisiva en la que se expone, describe o argumenta. Ámbitos público, educativo y profesional. Extensión: 550-600 palabras (máx. 5 minutos).
Tarea 5 6 ítems (2 ítems de B2 y 4 ítems de C1)	Comprender los puntos principales y extraer datos de los fragmentos de una conferencia.	Escuchar un monólogo dos veces y elegir, de entre tres opciones, una idea contenida en cada uno de los seis fragmentos en que está dividido (dos de ellos de nivel B2).	Conferencias, discursos, presentaciones y/o noticias radiofónicas o televisivas de extensión larga en los que se describen o narran proyectos y/o experiencias. Ámbito educativo y profesional. Extensión: 450-500 palabras (máx. 4 minutos).
Tarea 6 6 ítems (C1)	Comprender los puntos principales y extraer datos de una conferencia.	Escuchar un monólogo dos veces e identificar, entre doce opciones disponibles, seis ideas expresadas en el texto.	Conferencias, discursos, presentaciones y/o noticias radiofónicas o televisivas de extensión larga en los que se describen o narran proyectos y/o experiencias. Ámbito educativo y profesional. Extensión: 450-500 palabras (máx. 5 minutos).

PRUEBA 3. EXPRESIÓN E INTERACCIÓN ESCRITAS			
50 min 2 tareas	FOCALIZACIÓN Capacidad que se evalúa	FORMATO Tarea que se realiza	MATERIAL PARA LA TAREA
Tarea 1 Descripción y narración (A1-B1)	Informar, describir y narrar en un texto sencillo y cohesionado.	Redactar una respuesta a partir de una carta o un mensaje de correo electrónico, foro o blog.  El candidato lee un texto breve y sencillo, y redacta una respuesta de entre 100 y 150 palabras (recomendadas).	Un texto escrito de 30 a 50 palabras en forma de carta o mensaje (correo electrónico, foro o blog). Ámbito personal y/o público. Pautas para redactar el texto de salida, que propicien la producción de secuencias informativas, descriptivas y narrativas.
Tarea 2 Argumentación (B2-C1)	Exponer de manera clara, detallada y bien estructurada ideas, argumentos, opiniones y valoraciones, respetando las convenciones y rasgos del género textual al que pertenezcan.	Redactar un artículo de opinión o una carta a la redacción de un periódico a partir de un encabezado.  El candidato elige de entre dos opciones, lee un encabezado periodístico y redacta un texto de entre 250 y 300 palabras (recomendadas).	Un texto escrito de 50 a 70 palabras en forma de encabezado periodístico. Ámbito público, educativo y/o profesional. Pautas para redactar el texto de salida, que propicien la producción de un texto de carácter argumentativo.

PRUEBA 4. EXPRESIÓN E INTERACCIÓN ORALES			
De 15 a 20 minutos 5 tareas	FOCALIZACIÓN Capacidad que se evalúa	FORMATO Tarea que se realiza	MATERIAL PARA LA TAREA
Tarea 1 Preguntas personales (A1-A2) 1 minuto y 30 segundos	Proporcionar información sobre sí mismo, sus experiencias y su entorno más inmediato.	Responder a cuatro preguntas de carácter personal.	Cuatro preguntas de carácter personal: las dos primeras de nivel A1 y las dos últimas de nivel A2. Ámbito personal.
Tarea 2 Descripción (A2) 1-2 minutos + 1 minuto y 30 segundos de preparación	Describir, de manera breve y sencilla, los elementos de una escena de la vida cotidiana.	Desarrollar un monólogo breve a partir de una fotografía. El candidato elige entre dos fotografías, prepara y desarrolla la descripción de una de ellas.	Dos fotografías que reflejen una situación de la vida cotidiana. Ámbitos personal y/o público. Pautas sobre lo que ha de incluir la descripción.
Tarea 3 Simulaciones (B1) 2-3 minutos	Solicitar y dar información sobre necesidades inmediatas, mediante transacciones o contactos sociales.	Responder en dos situaciones simuladas: una del ámbito personal y otra del ámbito público. El candidato elige cada una de las situaciones de entre dos opciones.	Cuatro situaciones simuladas: dos pertenecientes al ámbito personal y dos al ámbito público. Ámbitos personal y/o público. Pautas para desarrollar la tarea.
Tarea 4 Preguntas (B2) 2-3 minutos	Expresar opinión, suposición y dar información respondiendo a una serie de preguntas sobre un determinado tema.	Responder a tres preguntas sobre un tema. El candidato elige un tema de entre dos opciones, lee un texto relacionado con ese tema y responde a tres preguntas.	Dos textos de diferente temática (80 a 120 palabras cada uno). Ámbitos público, educativo y/o profesional. Tres preguntas para cada tema.
Tarea 5 Monólogo (C1) 3-4 minutos + 2 minutos de preparación	Argumentar a favor o en contra de una afirmación sobre un determinado tema.	Desarrollar un monólogo. El candidato elige una afirmación de entre dos opciones, relacionadas con el tema de la tarea 4. Prepara y desarrolla la tarea.	Dos afirmaciones para cada tema. Ámbitos público, educativo y/o profesional. Pautas para desarrollar la tarea.

Como se puede apreciar, en todas las pruebas hay una tarea que corresponde con un nivel del MCER, excepto en el caso de la prueba de Expresión e interacción escritas, donde se consideró que era posible diferenciar entre los cinco niveles de desempeño evaluados a través de las escalas de calificación, con solo dos tareas.

Como hemos mencionado previamente, el examen SIELE nace con una clara vocación panhispánica y en él tienen cabida todas las variedades del español. En las tareas de niveles más bajos (A1-B1) se ha optado por una variedad neutra, sin rasgos dialectales o exclusivos de ninguna variedad, a la que se ha dado la etiqueta de “español internacional”.

En las pruebas de comprensión, los candidatos del SIELE se enfrentan a textos de al menos tres variedades distintas por prueba, lo que garantiza la correcta comprensión general por parte de los candidatos. En las pruebas de expresión, se aceptará cualquier variedad del español y se aplicarán los mismos criterios de calificación sea cual sea la escogida por el candidato.

### 3. Segunda fase de diseño: implementación informática

Tras la experimentación en papel, se inició la fase de diseño interactivo. Aunque en el diseño inicial se había tenido en cuenta que el examen se realizaría con ordenador, una vez que el equipo de Telefónica Educación Digital empezó la implementación de las tareas tuvieron que realizarse ajustes para que pudieran visualizarse correctamente en pantalla y para facilitar la interacción y navegación de los usuarios.<sup>2</sup>

Un aspecto fundamental en el diseño interactivo fue el equilibrio entre sencillez y simplicidad con la necesaria variedad en la tipología de tareas para contemplar diferentes tipos de candidatos y evitar un sesgo o inclinación favorable a cierto tipo de candidatos. En las pruebas del SIELE encontramos tareas de opción múltiple (con variaciones en el número de distractores y extensión de los textos), de relacionar, de ordenar, de selección múltiple con un banco común de respuestas, de escritura en campos de texto y de grabación de respuestas.

El desarrollo de las pruebas de comprensión no fue muy problemático, ya que la mecánica de realización de las tareas es perfectamente trasladable desde la experiencia de los exámenes con lápiz y papel; de hecho, el diseño interactivo facilita ciertos aspectos para los candidatos, como la recogida de respuestas en el mismo espacio en el que están las preguntas y no con una hoja preparada para la lectura óptica y ajena a la tarea, como generalmente sucede con los exámenes tipo DELE. La dificultad mayor en este tipo de pruebas residió en la necesidad de combinar en una misma pantalla los textos, las preguntas y los distractores de cada tarea, debido fundamentalmente a la gran extensión de los textos de las tareas de los niveles B2 y C1. Un aspecto que también cambia en el SIELE con respecto a los exámenes en soporte tradicional es que la navegación en estas pruebas es lineal de tarea en tarea: una vez que se han enviado las respuestas de una tarea, no se puede volver sobre lo realizado. Sin embargo, la navegación entre preguntas de una misma tarea sí es posible en la prueba de Comprensión de lectura. En la prueba de Comprensión auditiva la navegación está supeditada a las audiciones y el número de escuchas: las tareas se van sucediendo de manera automática, de acuerdo con las locuciones de los textos, de manera que los tiempos de respuesta y las condiciones de escucha son los mismos para todos los candidatos.

Sin duda, la prueba que resultó más fácilmente trasladable al contexto digital fue la prueba de Expresión e interacción escritas, no solo por la simplicidad técnica de su diseño interactivo (simplemente un cuadro de texto en el que escribir), sino también porque todos estamos familiarizados con la escritura en ordenador: de hecho, podemos decir que

---

<sup>2</sup> En la página <http://siele.org> se puede acceder a un modelo completo de examen para realizarlo en línea así como a la *Guía para la realización del examen SIELE*, donde se explica el funcionamiento de cada prueba.

las tareas de expresión escrita del SIELE resultan más auténticas que las que se realizan en los exámenes tradicionales con lápiz y papel. En esta prueba, sí es posible navegar entre las tareas, ya que se consideró que la revisión de lo escrito forma parte del constructo de la prueba y que los candidatos deben tener la autonomía necesaria para llevarla a cabo cuando ellos consideren oportuno. Además, al haber solo dos tareas de nivel muy distinto y estar destinadas a candidatos con diferentes grados de dominio, la decisión de cuánto tiempo se debe dedicar a cada una constituye una cuestión estratégica que puede condicionar los resultados. Otras utilidades que se incorporaron a la prueba son el contador de palabras, el acceso a los caracteres específicos del español mediante un teclado virtual y el control del tiempo.

Por supuesto, el mayor reto nos esperaba en el diseño de la prueba de Expresión e interacción orales, específicamente en las tareas de interacción. Resultaba evidente que en un examen de estas características no podíamos replicar la situación que se crea en las pruebas con entrevistador y que, en las tareas de interacción, el ordenador no podría interactuar con cada candidato de manera diferente a partir de sus respuestas.

Conscientes de ello, ya en la primera fase de diseño se buscó que las tareas para medir la interacción oral resultaran lo más naturales y auténticas que fuera posible en un contexto de grabación, seleccionando aquellas que realizamos con máquinas cada vez más frecuentemente, como por ejemplo, contestar a preguntas, o a otro tipo de mensajes grabados en contestadores o en aplicaciones como *Whatsapp*, y que pueden ser el primer momento de un intercambio comunicativo. Otro tipo de situaciones de interacción, como conversaciones informales, no aparecen en el examen y, por lo tanto, no se recogen en los descriptores que aparecen en los certificados para explicar el grado de dominio de los candidatos de acuerdo con su puntuación.<sup>3</sup>

---

<sup>3</sup> Se puede consultar el sistema de calificación y los descriptores de cada banda en el Capítulo 7 de la *Guía oficial para preparadores del SIELE* (en prensa).


**Expresión e interacción orales**

PUNTUACIÓN	NIVEL DEL MCER	DESCRIPCIÓN
215 - 250	C1	Los usuarios de la lengua de nivel C1 realizan descripciones y presentaciones claras y detalladas sobre temas complejos, integrando aspectos complementarios, desarrollando ideas concretas y terminando con una conclusión adecuada. Se expresan con fluidez y espontaneidad, casi sin esfuerzo, y dominan un amplio repertorio léxico.
176 - 214.99	B2	Los usuarios de la lengua de nivel B2 realizan descripciones y presentaciones claras y detalladas sobre una amplia serie de temas generales, académicos, profesionales o de ocio, desarrollando y defendiendo sus ideas con explicaciones y argumentos adecuados. Hablan con fluidez, precisión, control gramatical y eficacia, marcando con claridad la relación entre las ideas.
125 - 175.99	B1	Los usuarios de la lengua de nivel B1 son capaces de enfrentarse a la mayoría de las situaciones de un viaje, explicar un problema y hablar de asuntos habituales o de su interés y de temas relativamente abstractos y de carácter cultural (películas, libros, música, etc.). Se expresan con fluidez y seguridad razonables mediante un repertorio lingüístico sencillo que presentan como una secuencia lineal de elementos.
76 - 124.99	A2	Los usuarios de la lengua de nivel A2 se comunican en tareas sencillas y habituales que requieren un intercambio básico y directo de información. Son capaces de expresar opiniones personales y de hacer descripciones y presentaciones sencillas de personas, condiciones de vida, actividades diarias y cosas que les gustan, mediante breves listas de oraciones.
40 - 75.99	A1	Los usuarios de la lengua de nivel A1 son capaces de plantear y contestar preguntas habituales de áreas de necesidad inmediata o sobre temas muy cotidianos, así como de realizar afirmaciones básicas expresándose con frases sencillas y aisladas relativas a personas y lugares.
0 - 39.99	Inferior a A1	No hay descriptores de uso de la lengua por debajo del nivel A1.

En el diseño de esta prueba ha sido fundamental la automatización en la administración de las tareas, para que se sucedieran de manera natural y efectiva. En este sentido, la prueba oral del SIELE ofrece una gran fiabilidad o consistencia en las circunstancias de administración ya que se ha conseguido limitar las variaciones en la forma de aplicarse. Tras las experimentaciones que se han llevado a cabo, hemos visto con satisfacción que las muestras grabadas por los candidatos permiten evaluar de forma precisa la capacidad de los candidatos para comunicarse oralmente en relación con los niveles del MCER.

Tanto en la prueba de Expresión e interacción orales como en la de Expresión e interacción escritas, al medir destrezas productivas y ser de calificación subjetiva, el proceso de diseño implica un paso más: el desarrollo de escalas de calificación. En el caso de las escalas del SIELE se ha hecho un esfuerzo muy importante para simplificar y ajustar el proceso de calificación y garantizar la fiabilidad.

Las escalas del SIELE se han elaborado a partir de las tablas del MCER, los inventarios del PCIC y las características de las tareas. Las escalas son tablas con:

- un eje vertical en el que se distinguen seis bandas (0-5), que reflejan los distintos niveles de respuesta;
- y un eje horizontal con las categorías correspondientes a los aspectos que se valoran en la calificación.

En cada categoría hay, para cada banda, un descriptor que define un determinado nivel de desempeño. El calificador debe elegir la banda en la que mejor se describe la respuesta del candidato.

Las escalas de la prueba de EIE tienen un carácter más analítico, con tres categorías lingüísticas (Cohesión, Corrección y Alcance) y una de Cumplimiento de la tarea, que es diferente para la tarea 1 y para la tarea 2. Las seis bandas de las categorías lingüísticas se corresponden con los niveles del MCER pero, lógicamente, las de cumplimiento, no. Hay que señalar que existe una cierta ponderación hacia el cumplimiento, en consonancia con el constructo de la prueba, basado en un enfoque comunicativo y orientado a la acción.

Antes de describir las escalas, es necesario señalar que tanto en la prueba de EIO como en la prueba de EIE, la calificación la lleva a cabo un profesional de la enseñanza de español como lengua segunda o extranjera que ha sido acreditado como calificador SIELE tras realizar y superar satisfactoriamente un curso de formación específico.

En la prueba de EIO, se ha optado, sin embargo, por dos categorías, Uso de la lengua y Cumplimiento de la tarea, que se distribuyen en dos escalas: una para las tareas 1, 2 y 3 y otra para la 4 y la 5. La reducción en el número de categorías en esta prueba se debe a que algunas de las cinco tareas de la prueba son tan breves que no es pertinente aplicar todos los criterios que se empleaban para las dos tareas de la prueba escrita. De esta forma, tanto los rasgos que en la prueba de EIE se dividían en tres categorías lingüísticas como otros característicos del lenguaje oral como, por ejemplo, la fluidez y la pronunciación, se agrupan en Uso de la lengua.

Una vez completado el diseño, se crearon tres formas de examen que se experimentaron con 200 candidatos cada una. De forma paralela se llevó a cabo un juicio de expertos para determinar la dificultad de las tareas de comprensión y establecer los puntos de corte entre los niveles. En el juicio de expertos, participaron treinta profesores de la UNAM, la USAL y del Instituto Cervantes.

En las pruebas de expresión, se llevaron a cabo dos talleres de estandarización y calibración de las muestras obtenidas en las experimentaciones, en los que se revisaron las escalas de calificación, se ajustaron a las tareas y se establecieron los puntos de corte de estas pruebas.

#### **4. Proceso de creación de tareas y pruebas**

Una de las aportaciones fundamentales del SIELE es que se ha conseguido desarrollar un proceso de creación de tareas común y compartido entre las instituciones participantes que garantiza la homogeneidad, calidad y validez de todas las formas de examen desarrolladas.

En primer lugar, cada institución crea un conjunto de tareas que debe pasar por unos filtros internos. Para esta fase de creación, se han desarrollado un conjunto de herramientas comunes que manejamos todos los equipos:

- Plantillas para el desarrollo de cada tarea.
- Fichas para su validación basadas en las especificaciones.
- Ficheros comunes con los históricos de todas las tareas creadas hasta el momento.
- Manual de estilo del SIELE.

Hay que señalar que, tal y como indicamos al principio de este artículo, se tomó la decisión de utilizar los inventarios del PCIC como referencia en cuanto a los contenidos lingüísticos de cada tarea y esto se está realizando de manera rigurosa.

Una vez validadas en cada institución, las tareas se envían a otra institución que debe revisarlas haciendo uso de las herramientas comunes, comentando todos los aspectos mejorables y proponiendo los cambios pertinentes.

Estas tareas revisadas pasan a una tercera institución que debe corroborar las propuestas de la segunda o de la primera institución.

Al finalizar este primer ciclo de revisión, la institución creadora debe actualizar sus propuestas teniendo en cuenta los comentarios recibidos. Una vez introducidos los cambios y mejoras, el ciclo de revisión se vuelve a iniciar para así validar definitivamente las tareas.

Creemos que este doble proceso de revisión por parte de al menos dos instituciones ha contribuido notablemente al desarrollo de una cultura común de evaluación. Del mismo modo, la creación compartida y el análisis minucioso de las tareas facilita la integración de las diferentes variedades lingüísticas en las pruebas y su adecuado uso en las tareas de los diferentes niveles.

## BIBLIOGRAFÍA

- CONSEJO DE EUROPA (2001), *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Madrid, Instituto Cervantes -Ministerio de Educación, Cultura y Deporte- Anaya [en línea] <<http://cvc.cervantes.es/obref/marco>> [consulta: 14/12/2016]
- CONSEJO DE EUROPA (2009), *Manual for Relating Examinations to the Common European Framework of Reference for Languages*. Estrasburgo, Language Policy Division [en línea] <[http://www.coe.int/t/dg4/linguistic/Manuel1\\_EN.asp#Manual](http://www.coe.int/t/dg4/linguistic/Manuel1_EN.asp#Manual)> [consulta: 14/12/2016]
- INSTITUTO CERVANTES (2006), *Plan curricular del Instituto Cervantes. Niveles de referencia para el español*. Madrid, Biblioteca Nueva.
- SIELE (2016), *Guía para la realización del examen SIELE*. Madrid, Telefónica Educación Digital [en línea] <<https://siele.org/examen>> [consulta: 14/12/2016]

SIELE (2017), *Guía oficial para preparadores del SIELE*. Madrid, Telefónica Educación Digital.