



4^{to} Encuentro internacional de español como lengua extranjera

Enseñanza, aprendizaje y evaluación

LA VIABILIDAD, EL ENFOQUE COMUNICATIVO Y LA CLASIFICACIÓN EN LA EVALUACIÓN DE LA LENGUA ORAL

Javier Hernán Bejarano Ruiz¹ | bejarano.javier@javeriana.edu.co
Pontificia Universidad Javeriana

Introducción

Este artículo surge como parte del trabajo final de investigación de la maestría en Lingüística Aplicada del Español como Lengua Extranjera (ELE) de la Pontificia Universidad Javeriana (PUJ) (2013-2015). El proyecto investigativo tenía como objetivo diseñar una propuesta de prueba oral de clasificación para el Centro Latinoamericano (CLAM) de la PUJ.

Son muy diversas las motivaciones que tienen las personas para estudiar una lengua extranjera. Por ejemplo, hacer un proceso de inmigración, conseguir un nuevo trabajo, realizar estudios en el exterior, certificar un nivel de dominio de la lengua en un examen internacional, ampliar el conocimiento del mundo o, incluso, entender algunos aspectos de otra cultura, entre otras. Antes de lograr tal propósito y obtener una certificación que reconozca dicho esfuerzo, el primer paso, en cualquier institución de enseñanza de lenguas extranjeras, es presentar un examen de clasificación².

Al igual que otros centros en Colombia y en el mundo, el CLAM utiliza un examen de clasificación o de nivel para clasificar a los nuevos estudiantes que estén interesados en ingresar al programa de ELE. A pesar de que el actual examen tiene un rango mínimo de error al momento de clasificar a un estudiante, se evidenció que la prueba de expresión oral tenía una discrepancia entre el cómo se enseña, un enfoque comunicativo fuerte (enseñanza mediante tareas) y el cómo se evalúa, un enfoque comunicativo débil (formato de preguntas para la

¹ Magíster en lingüística aplicada del español como lengua extranjera y licenciado en lenguas modernas de la Pontificia Universidad Javeriana. Profesor de cátedra de francés y español de la Pontificia Universidad Javeriana, Departamento de Lenguas, Facultad de Comunicación y Lenguaje.

² Para efectos de este artículo, se llama examen a un método de evaluación que incluye diferentes secciones o pruebas (prueba de comprensión oral, de producción escrita, etc.).

evaluación lingüística). No obstante, aunque se hizo una propuesta de prueba oral de clasificación que intentara responder a la enseñanza mediante tareas, se demostró que ésta no era viable para la clasificación. Por lo anterior, en la clasificación, la evaluación no necesariamente puede estar en relación con el programa de enseñanza.

Por otro lado, a lo largo de esta investigación se pudo constatar que el campo de investigación en evaluación ha sido poco abordado y tratado en la enseñanza de lenguas extranjeras y particularmente en español y en Colombia. La literatura que se encuentra está escrita en inglés y se centra más en los principios y el diseño de un examen de dominio como el IELTS en inglés o el DELE en español. Es poco lo que se ha escrito sobre la clasificación a pesar de que se usa con regularidad en las instituciones. Por esta razón, este artículo puede contribuir a ampliar esta literatura investigativa y especialmente en la evaluación y clasificación de la lengua oral.

Prueba de clasificación

Para comenzar, de manera general, un prueba es un método que incluye técnicas, procedimientos e ítems para medir la habilidad, el conocimiento o la actuación de una persona (Brown, 2005). Bachman y Palmer (1996) afirman que el hecho de diseñar o aplicar un prueba no implica que ésta sea útil. Por esta razón, los autores proponen un modelo de seis cualidades para evaluar no solamente el diseño y el uso de un prueba, sino también todos los aspectos de su diseño y su aplicación. Estas cualidades son: validez, fiabilidad, autenticidad, interactividad, impacto y viabilidad.

A *grosso modo* la validez determina hasta qué punto una prueba mide lo que supone que mide en relación con su propósito (Mingarro, 2008). Una prueba no es válida solo asegurando que lo es. Por lo tanto, su validez debe establecerse y demostrarse de diferentes formas. Alderson et al (1995) engloban la validez en tres grandes tipos: validez interna, validez externa y validez del constructo.

La validez interna se refiere al análisis del contenido mismo de la prueba y su efecto, y la validez externa se refiere a cualquier otro tipo de información que demuestre una relación entre los resultados de la prueba y algún otro criterio que se crea que es también un indicador de la habilidad evaluada. En lo que concierne la validez del constructo, ésta hace referencia a lo que se mide en la prueba y sirve para determinar si las partes de la prueba miden lo que deben medir y no otras cosas (Martínez, 2008).

En términos sencillos la fiabilidad es la medida en que los resultados de un instrumento de medición son consistentes en la repetición. Es decir, si un candidato presentara la misma prueba varias veces, los resultados deberían ser los mismos (Mingarro, 2008). La fiabilidad tiene como objetivo minimizar los efectos de error de medición. Es decir, una prueba es fiable cuando logra disminuir el efecto de varios factores, como la falta de motivación o interés, la fatiga y otras variables externas que puedan afectar los resultados (Bachman, 1990). La fiabilidad, al igual que la validez, puede estimarse a partir de diferentes procedimientos como el test-retest, la fiabilidad paralela y el método de las dos mitades entre otros.

Para Bachman y Palmer la autenticidad es el grado de correspondencia entre las características de las actividades³ en la prueba y las características de las tareas en el uso de la lengua en la vida real. Dicho de otro modo, hasta qué punto las tareas utilizadas en la prueba para evaluar la actuación de un candidato se relacionan con lo que solemos hacer en nuestra vida cotidiana cuando hacemos uso de la lengua y el lenguaje. La autenticidad es importante no solo porque puede establecer una relación entre las tareas de la prueba y la vida diaria, sino también porque éstas pueden influenciar la actuación del candidato en la misma prueba.

La interactividad corresponde al grado en que las actividades evaluativas de la prueba involucran la competencia comunicativa⁴, el conocimiento del tema y el esquema afectivo del candidato. En otras palabras, hasta qué punto las actividades llevan al candidato no solamente a utilizar la lengua y el lenguaje, sino a involucrarlo afectivamente según el tema para poder resolver el objetivo de la tarea.

El impacto hace referencia al grado en el que los resultados afectan a los individuos, las instituciones y a la sociedad en general. Los resultados de una prueba afectan positiva o negativamente a los participantes que lo presentan y la institución que lo aplica. Por ejemplo, la familiaridad con las técnicas e instrucciones de una determinada prueba afectarán a los individuos que lo presenten.

Finalmente, la viabilidad corresponde a la relación existente entre los recursos requeridos para el diseño y desarrollo de la prueba, y los recursos que en efecto estarán disponibles para esos

³ En la didáctica del aprendizaje de lenguas los ejercicios han sido rebautizados como actividades o tareas. La tarea tiene una connotación más real mientras las actividades son simulaciones de situaciones (Vázquez, 2000).

⁴ Varios autores han definido este concepto. Una definición muy sencilla es la de Hymes (1967): es la capacidad de saber qué decir a quién y cómo decirlo de la forma más apropiada en una situación determinada.

efectos (Bachman y Palmer, 1996). En otras palabras, cuáles son los recursos humanos, los materiales, el espacio y el tiempo necesario para el diseño, el pilotaje, la evaluación, la administración de la prueba, la aplicación y la calificación de la prueba.

Puede decirse entonces que al momento de diseñar una prueba de clasificación debería tenerse en cuenta las anteriores cualidades para su proceso de diseño y evaluación. El propósito de una prueba de clasificación es clasificar al estudiante dentro de un nivel apropiado en un currículo de lengua (Brown, 2005). De igual forma, juega un papel importante ya que de su eficacia puede depender la cohesión y el buen funcionamiento del grupo, para evitar desequilibrios entre niveles de conocimiento y uso de la lengua (Pastor, 2003).

Por otro lado, una prueba de clasificación debería estar basada en el marco de un programa de enseñanza y en los objetivos y contenidos propuestos y enseñados por la institución que lo realice (Brown, 2005). Por esta razón, es la institución quien decide, según su programa de enseñanza y su contexto, qué niveles y habilidades necesitan evaluar (Martínez, 2011). Esta autonomía institucional puede evidenciarse en la variedad de instituciones de enseñanza de lenguas extranjeras, especialmente en inglés y francés, en Colombia. Existe una variedad de pruebas con diferentes secciones y que evalúan diferentes habilidades. Cabe preguntarse si dichas pruebas están en relación con los programas de la institución y qué cualidades tienen.

El enfoque comunicativo

Hasta el momento se podría decir que una prueba de clasificación debería no solamente estar relacionada con los objetivos y contenidos del programa de enseñanza, sino cumplir con las cualidades anteriormente mencionadas. Por lo anterior, en vista de que el CLAM tiene un programa de enseñanza mediante tareas en el marco del enfoque comunicativo, su evaluación debería seguir esta línea. Pero ¿qué es el enfoque comunicativo y la enseñanza mediante tareas?

Desde los años 70 con la aparición del concepto de competencia comunicativa de Hymes se dio una nueva dirección a la enseñanza de lenguas. El denominado enfoque comunicativo que defiende la concepción comunicativa de la lengua, traería una nueva perspectiva en oposición al enfoque estructuralista que defendía el uso correcto y estructural de la lengua (Font, 1998). En palabras de Hymes, la competencia comunicativa es la capacidad de formar enunciados que sean socialmente apropiados y gramaticalmente correctos (Instituto Cervantes, 1997-2016). Hymes afirma que conocer una lengua es más que conocer sus reglas gramaticales, pues también hay

reglas culturales específicas del uso de la lengua que se relacionan con características del contexto comunicativo (citado en McNamara, 2000).

Desde la aparición del concepto de competencia comunicativa, varios modelos (Canale y Swain en los 80, Bachman, Celce-Murcia, Dörnyei y Thurrell en los 90) se han propuesto para saber qué conocimientos, capacidades o destrezas se necesitan para hacer un uso comunicativo de la lengua. Para efectos de este trabajo se tuvo en cuenta la definición de competencia comunicativa definida por el MCER que la clasifica como el componente de tres subcompetencias: la competencia lingüística, la competencia sociolingüística y la competencia pragmática.

La competencia lingüística incluye los conocimientos y las destrezas léxicas, fonológicas y sintácticas, y otras dimensiones de la lengua como sistema, independientemente del valor sociolingüístico de sus variantes y de las funciones pragmáticas de sus realizaciones (Consejo de Europa, 2002: 13).

La competencia pragmática tiene que ver con el uso funcional de los recursos lingüísticos sobre la base de guiones o escenarios de intercambios comunicativos. También tiene que ver con el dominio del discurso, la cohesión y la coherencia, la identificación de tipos y formas de texto, la ironía y la parodia. (Consejo de Europa, 2002: 14).

La competencia sociolingüística se refiere a las condiciones socioculturales del uso de la lengua. Mediante su sensibilidad a las convenciones sociales (las normas de cortesía, las normas que ordenan las relaciones entre generaciones, sexos, clases y grupos sociales, la codificación lingüística de determinados rituales fundamentales para el funcionamiento de una comunidad) (Consejo de Europa, 2002: 13).

Esta visión de la competencia comunicativa implica un conjunto de conocimientos formales e instrumentales de la lengua extranjera, así como contenidos y procesos para la comunicación (Estaire y Zanón, 1990), en otras palabras, es saber qué decir a quién y cómo decirlo de forma apropiada en una situación y contexto determinado.

Aunque el enfoque comunicativo no determina el diseño (objetivos, contenidos, actividades, etc.) y el procedimiento (recursos en el aula, prácticas del docente, etc) de un único método de enseñanza y aprendizaje, ya que puede desarrollarse e interpretarse teniendo en cuenta la realidad en el aula y las necesidades de los estudiantes (Melero, 2008), Stern (citado en Estaire, 2010)

hace una división en dos grandes grupos: un enfoque comunicativo de tipo lingüístico (débil) y uno de tipo psicolingüístico (fuerte).

En un enfoque débil se diseñan actividades comunicativas a partir de los contenidos lingüísticos que se presentan y se practican siguiendo la secuencia de presentación, práctica y producción (PPP). Como por ejemplo, los programas nocio-funcionales⁵ que se han desarrollado desde los años 70. Por otro lado, un enfoque fuerte diseña un repertorio de tareas comunicativas que se ajustan y favorecen los procesos implicados en la adquisición de una lengua. Por ejemplo, la enseñanza mediante tareas⁶. ¿Pero qué se entiende por tareas?

Para Long (como se citó en Estaire, 2010: 4) las tareas se definen como “todo aquello que solemos hacer en nuestra vida diaria, en el trabajo, en el tiempo libre, o en cualquier otro momento”. Es decir, cualquier actividad donde utilicemos la lengua; por ejemplo, ir al cine, ir a un restaurante, comprar algo, escuchar y debatir un programa de televisión etc.

El enfoque comunicativo fuerte, como afirman Estaire y Zanón (1990), aboga por el desarrollo de la competencia comunicativa a partir de unidades, tareas comunicativas, que integren no solamente el conocimiento formal del lenguaje (los contenidos para la comunicación), sino también el conocimiento instrumental (capacidad de usar el lenguaje) para alcanzar un objetivo comunicativo. Por esta razón, Ellis (citado en Soler, 2008) afirma que una tarea comunicativa debería implicar el uso del lenguaje donde se emplean los mismos procesos comunicativos que en las actividades del mundo real para la transmisión de un mensaje.

El enfoque comunicativo fuerte para la evaluación

Hasta el momento se ha hablado de un enfoque comunicativo fuerte para la enseñanza y el aprendizaje. En lo que concierne la evaluación, Soler (2008) afirma que la evaluación también encontró argumentos en la teoría de la competencia comunicativa para servirse de tareas comunicativas para la evaluación. De igual forma, también hay un interés en el uso de tareas para provocar la actuación de un candidato y para realizar una evaluación con necesidades específicas, es decir, una evaluación, como afirma Ellis (citado en Soler, 2008: 113), “que utiliza tareas

⁵ Nacionales: conceptos que se expresan a través de la lengua como son tiempo, cantidad, frecuencia, ubicación y secuencia. Funcionales: intenciones del hablante como por ejemplo presentarse, pedir información en un cine, quejarse, etc.

⁶ En esta línea se puede también hablar de los programas de procesos y proyectos.

holísticas que implican conductas reales (o lo más cercano a éstas), que ponen en marcha el tipo de procesamiento lingüístico que se daría en actividades del mundo real”.

Por otro lado, siguiendo la línea de la enseñanza, puede decirse entonces que una evaluación puede hacerse bajo un enfoque fuerte o débil. El enfoque débil consiste en tareas que no necesariamente están relacionadas con el mundo real donde la actuación se juzga en función de la habilidad lingüística. En el enfoque fuerte las tareas están relacionadas con el mundo real y su actuación se juzga con criterios del mundo real, es decir, como afirman Long y Norris (citados en Soler, 2008), las tareas del mundo real no solo provocan el uso de la lengua, sino que se evalúa la realización de la tarea misma.

La evaluación de la lengua oral

La lengua oral puede evaluarse a partir de diferentes actividades según el programa de aprendizaje. Autores como Luoma (2004) hacen la distinción de dos tipos de actividades para evaluarla. Las actividades abiertas, que tienen como propósito conseguir que el candidato haga algo con la lengua como indicador que posee ciertas habilidades⁷, y las actividades cerradas, que especifican lo que el candidato debe decir para evaluar aspectos precisos de la lengua⁸. Así por ejemplo, una actividad abierta de *decisión*, llevan al candidato a argumentar en pro o en contra sobre una solución a un problema.

Por otro lado, una actividad cerrada de *leer en voz alta*, tiene como objetivo evaluar diferentes aspectos de la pronunciación y de la comprensión como el ritmo, el estrés, la entonación y la precisión de algunos sonidos (Luoma, 2004). Puede verse que una actividad abierta se acerca más a una evaluación de la competencia comunicativa ya que puede llevar al candidato no solamente a utilizar su conocimiento formal del lenguaje (competencia lingüística), sino también el uso del lenguaje (competencia pragmática, estratégica, discursiva y sociocultural) para alcanzar un objetivo comunicativo. Una actividad cerrada, por otro lado, lleva al candidato a utilizar solamente su conocimiento formal del lenguaje (gramática, pronunciación y vocabulario).

Una actividad abierta, una tarea, pueden realizarse entre: un candidato y un entrevistador, que hace también el papel de evaluador; entre un candidato, un entrevistador y un evaluador; o entre

⁷ Las actividades pueden ser de *descripción, narración, instrucción, comparación, explicación, justificación, predicción, decisión, juego de rol y la reacción a una situación.*

⁸ Las actividades pueden ser de *preguntas-respuestas, leer en voz alta, repetir una frase, completar una frase y reaccionar a una situación.*

dos estudiantes o más y un evaluador (Pastor, 2007). Entre estas formas de realizar la evaluación, Luoma (2004) afirma que la más común es la entrevista entre el candidato y el evaluador ya que es flexible, se adapta a la actuación del candidato y el evaluador puede tener control sobre la interacción, además ésta tiene que estar estructurada para que todos los aspectos de la prueba se cumplan y todos los candidatos sean evaluados igualmente (principio de fiabilidad).

En una entrevista, puede evidenciarse que aislar la comprensión auditiva en un intercambio entre un examinador y un candidato resulta bastante difícil. Bordon (2004) afirma que las instrucciones son orales éstas deben ser cortas y de reacción inmediata para que el candidato pueda recordarlas; de lo contrario, es recomendable que éstas estén escritas en una ficha de simulación o como parte de la tarea para facilitar la comprensión de lo que el candidato debe hacer. En lo que concierne las instrucciones para el evaluador o entrevistador, éstas deben estar por escrito, ser claras y explicar lo qué se debe hacer y cuándo hacerlo (Alderson et al, 1995).

En lo que concierne a la cantidad de tareas que deberían utilizarse, Bordon (2004) dice que es recomendable utilizar dos o tres tareas, pero sin abusar de la variedad y empezando por la tarea más sencilla y hasta la tarea más compleja y larga (aprox. 5 minutos). Alderson et al (1995) afirman que no existe una única y mejor forma para evaluar una habilidad; por consiguiente, recomienda que se utilice más de una forma para que pueda estar dirigida a diferentes tipos de candidatos.

Con respecto al tiempo Bordon (2004) dice que el tiempo debe ser suficiente para obtener una muestra valiosa de la actuación del candidato y recomienda mínimo 10 minutos y máximo 30 minutos. Por su parte, Luoma (2004) afirma que el tiempo lo decidirán los recursos disponibles que se tienen, la viabilidad. Por otra parte, las pruebas orales de los exámenes de dominio internacionales, pueden servir de referente para considerar cuánto tiempo podría durar una prueba oral. Por ejemplo, en español, en los exámenes DELE A1, A2, B1 y CELU la prueba oral tienen una duración de 15 minutos, y en inglés, en el examen IELTS, una duración de 11 a 14 minutos.

Diseño de la prueba

Cuando se pretende desarrollar un examen, o en este caso una prueba, casi siempre se tiene una idea general de lo que se va a desarrollar. Normalmente se sabe o se intuye cuál es la visión de lengua que se tiene, cómo se va a evaluar, qué criterios se van a considerar y en qué tipo de programa podría enmarcarse (Martínez, 2011). Sin embargo, para llevar este conjunto de ideas a

una descripción detallada y un plan de acción, debe recurrirse a un modelo de diseño y de elaboración.

Existen diferentes modelos que explican la estructura y el propósito de un examen. Alderson et al (1995) afirman que un examen debería tener:

- ♦ Un propósito
- ♦ Una descripción del candidato
- ♦ Una definición del constructo
- ♦ Diferentes secciones con su respectiva ponderación de cada sección
- ♦ Situaciones contempladas en la lengua meta
- ♦ Tipos de textos
- ♦ Tipos de tareas
- ♦ Instrucciones
- ♦ Criterios de evaluación
- ♦ Descripciones lingüísticas para cada nivel
- ♦ Descripciones de lo que los candidatos pueden hacer en el mundo real

Otros autores como Luoma (2004) y McNamara (2000) también consideran estas características para la elaboración de un examen; sin embargo, para este proyecto se decidió adoptar el modelo de Bachman y Palmer (2010). Este modelo propone cinco etapas: 1) la planificación, 2) el diseño, 3) el plan de acción, 4) la prueba experimental y 5) el uso del examen.

La planificación consiste en hacer un análisis de necesidades y de alcance para sustentar que existe la necesidad de diseñar un examen y determinar cuáles son los recursos que se tienen para tal propósito. Para este proyecto de investigación, como ya se mencionó, existe la necesidad de hacer una propuesta de prueba oral que se articule al programa de enseñanza. En lo que concierne el análisis del alcance, se determinó que se haría el diseño, pilotaje y validez de tres (3) tareas y de los criterios de evaluación. En la etapa del diseño, primero se hace una descripción de los participantes involucrados en la prueba. En este caso, se hizo una descripción del tipo de estudiante (adultos extranjeros de diferentes nacionalidades y culturas) y de los evaluadores (profesores del CLAM con experiencia en evaluación y enseñanza de ELE).

Por otra parte, se hace una definición del constructo a ser evaluado. Es decir, los desarrolladores deciden acerca de las habilidades que se van a evaluar. En el caso de esta

propuesta, se consideraron los aspectos lingüísticos de: léxico, gramática y pronunciación, y pragmáticos de: interacción, fluidez y coherencia. Estos aspectos son los mismos que considera el MCER en los aspectos cualitativos de la lengua (Consejo de Europa, 2002) y que pone a disposición, entre las 14 posibles escalas para evaluar la lengua oral (Consejo de Europa, 2002). De igual forma, estos criterios los utilizan los exámenes internacionales DELE en español y IELTS en Inglés, lo que es un referente fiable, por su experiencia y reconocimiento, para evaluar la lengua oral.

Luego de definir el constructo, se realiza una descripción detallada del nivel de dominio de la lengua. Para estos efectos se crearon descriptores detallados y específicos (analíticos) que dan cuenta de cada nivel de lengua correspondiente a los diferentes cursos del CLAM. El proceso de creación de dichos descriptores consistió en un análisis y paralelo entre: a) la escala cualitativa de la lengua propuesta por el MCER para todos los niveles, b) los criterios de evaluación de la banda dos (2) propuestos por el DELE para todos los niveles, c) los objetivos generales del Plan Curricular Instituto Cervantes (PCIC), d) el inventario de gramática y las nociones específicas propuestas por el PCIC y e) los objetivos generales y los programas de cada nivel del CLAM.

CRITERIO NIVEL	ALCANCE	CORRECCIÓN	FLUIDEZ	INTERACCIÓN	COHERENCIA
B1.1	Tiene un repertorio lingüístico amplio para desenvolverse, a pesar de algunos circunloquios, en situaciones cotidianas de temas como su familia, sus hábitos, su trabajo, su alimentación, sus pertenencias, sus aficiones e intereses. Puede contar experiencias en el pasado sobre viajes, tipo de viajes y medios de transporte	Muestra un control razonable de un repertorio de fórmulas y tiempos verbales de uso habitual como el presente, el pretérito perfecto, el pretérito imperfecto, el pretérito indefinido con irregularidades consonánticas, vocálicas y algunas perífrasis verbales como “seguir + gerundio” y “llevar + gerundio”. Comete algunos errores que provocan repeticiones y reformulaciones	Tiene un discurso continuado y comprensible aunque son evidentes sus pausas para planificar su discurso y para pensar en la gramática y el léxico adecuados. Se expresa un tanto dubitativamente y con algunos circunloquios sobre todo cuando aborda temas poco familiares	Inicia, mantiene y termina conversaciones cara a cara sencillas sobre temas cotidianos o que le son familiares. En ocasiones necesita pedir aclaración o repetir parte de lo que le han dicho para confirmar la comprensión mutua.	Enlaza una serie de elementos breves, diferenciados y sencillos con conectores como “primero”, “después”, “de pronto”, “de repente”, “al final” para formar una secuencia lineal de ideas relacionadas

Imagen 1: Criterios y descriptores para el cuarto (4) nivel del programa de enseñanza del CLAM correspondiente a un nivel B1.1

Finalmente, después de haberse definido el constructo a evaluar y hacer una descripción detallada del nivel de dominio de la lengua, se hace la construcción y selección de las tareas evaluativas. Una vez seleccionadas las tareas de lengua, es necesaria la descripción detallada de las características de éstas. Es decir, el tipo de tarea (abierta o cerrada), su función, el nivel a evaluar, y su técnica (descripción, narración, etc.).

En Bogotá existen diferentes club de conversación. Cada semana se escogen temas interesantes para que extranjeros practiquen el español. Hoy se va a hablar de viajes.

- ✓ Imagina que el entrevistador y tú son parte de este club
- ✓ Habla acerca de un viaje interesante que hiciste
- ✓ Guíate por las siguientes preguntas:

- ♦ ¿Dónde estviste?
- ♦ ¿Cuánto tiempo estuviste allí?
- ♦ ¿Con quién fuiste?
- ♦ ¿Qué lugar te gustó más? ¿Qué había allí?

Imagen 2: Tarea B para evaluar cuatro (4) niveles del programa de enseñanza.

Esta tarea tiene como función llevar al candidato, primero, a contar experiencias en el pasado sobre viajes, segundo, a describir lugares y personas (ver los criterios del ejemplo anterior), y tercero, a tener una conversación con el entrevistador donde expresa sus opiniones sobre el mismo tema. La técnica utilizada es el juego de rol donde se simula una situación de comunicación entre el candidato y el entrevistador. El candidato asume el papel de un estudiante que empieza una conversación sobre viajes en un club de conversación. A su vez, el entrevistador juega el rol de otro estudiante que promueve la conversación sobre el mismo tema.

El plan de acción tiene como objetivo principal guiar el proceso de diseño y aplicación de las tareas evaluativas. Éste consiste en una descripción detallada de la estructura del examen (especificaciones de la prueba), las especificaciones de cada tarea, y los procedimientos para administrar la prueba.

Las especificaciones de la prueba son todas aquellas características que prescriben la estructura y organización de ésta. Es decir, el número de partes que la conforman y su secuencia, el número y el valor de las tareas evaluativas, y el tiempo acordado para cada tarea y para el total de la prueba. Para esta propuesta solo se hizo una sección, la prueba de producción oral, que hará parte de las otras secciones del examen total. Se tomó la decisión de hacer 3 tareas evaluativas

para evaluar los 12 niveles del CLAM. Así, cada tarea evalúa 4 niveles⁹. Respecto a la secuencia de las tareas, se parte de menor a mayor nivel, puesto que es una prueba de clasificación. Finalmente se determinó que el total de la prueba sería de 20 minutos, 5 minutos para las tareas A y B, 7 minutos para la tarea C y 3 minutos para la calificación por parte del evaluador.

En lo que concierne las especificaciones de cada tarea, vuelven y se presentan las tareas evaluativas y los criterios y descriptores que den cuenta del constructo a ser evaluado, pero esta vez para el evaluador. Finalmente, con respecto a los procedimientos para administrar la prueba, se toman decisiones logísticas concernientes al espacio (edificio, salones, oficinas, etc.), al tiempo (tiempo de la prueba, hora de llegada de los candidatos, etc.), a los recursos humanos (el personal de recepción, los evaluadores, los asistentes, etc.) y a los procesos administrativos (inscripción de candidatos, verificación de la identidad, etc.).

Resultados

Durante la etapa de la prueba experimental los diseñadores del examen podrán recoger diferentes tipos de retroalimentación que eventualmente servirán para hacer modificaciones a las tareas y para reflexionar acerca del uso del examen en un contexto real, especialmente en lo que concierne el uso de recursos como tiempo, espacio, profesores requeridos, etc. Esta retroalimentación puede hacerse a partir de cuestionarios, observaciones y descripciones, entrevistas, o análisis estadísticos, entre otros, y será el último paso antes de la versión final de la prueba.

Los resultados obtenidos en esta etapa provienen de los siguientes instrumentos y actividades de recolección de información:

- Cuestionarios de fiabilidad y de validez del constructo de Bachman y Palmer (1996) para un grupo de docentes y expertos en evaluación en ELE.
- Cuestionario de valoración del contenido de las tareas, las instrucciones, el material y el tiempo de cada tarea por docentes y expertos.
- Comentarios y sugerencias de los evaluadores con respecto a las tareas y los criterios y descriptores de evaluación durante la evaluación y el pilotaje de las tareas.

⁹ La tarea A (nivel A1 y A2) para los niveles de 1 a 4. La tarea B (nivel B1 y B2) para los niveles 5 a 8. Y la tarea C (nivel C1 y C1) para los niveles 9 a 12.

Los resultados llevaron, como en todo proceso de diseño de instrumentos de evaluación, a seguir en la construcción de los criterios y descriptores que definen el constructo de lengua a evaluar. Se afirmó que el criterio de *Corrección* no solo debería hacer referencia a la competencia gramatical del hablante, sino también a la corrección de errores por medio de estrategias como la petición de aclaración, el ofrecimiento de aclaración y la reparación de la comunicación (Consejo de Europa, 2002). De igual forma, para el criterio de *Pronunciación*, se estableció que no solamente se trata de la habilidad de producción de sonidos, rasgos fonéticos y prosodia, sino también de la habilidad de percibirlos. Finalmente, aunque la coherencia se definió como la capacidad del hablante para organizar fragmentos coherentes de lengua, los descriptores se enfocan más a evaluar la cohesión¹⁰.

Por otro lado, aunque los criterios escogidos describen en general el constructo de lengua a evaluar, se propuso intentar evaluar otros aspectos de la lengua como los aspectos sociolingüísticos y paralingüísticos. Es decir, la evaluación no debería limitarse a la evaluación de la competencia lingüística y pragmática, sino intentar evaluar elementos que también forman parte de la competencia comunicativa como son los aspectos no verbales (componente estratégico) y las relaciones entre los hablantes.

Respecto a los descriptores de los criterios para cada uno de los niveles, se evidencian opiniones diferentes de los expertos. Por un lado, algunos afirman que el uso de descriptores detallados y específicos es más apropiado para la clasificación de la actuación de un candidato, ya que hay una descripción detallada de la progresión del estudiante en cada uno de los aspectos a evaluar. Por otro lado, otros expertos, en función de la viabilidad, afirman que los descriptores podrían resumirse en puntos clave para que sea visualmente más fácil calificar al candidato, pues la lectura detallada puede hacer que no se tenga el tiempo necesario para evaluar cada aspecto de la lengua durante y después de la actuación del candidato.

En lo que conciernen los resultados de las tareas, se determinó que éstas no son las definitivas y se hacen las siguientes sugerencias. La tarea A no necesariamente lleva al uso del lenguaje en el mundo real. Se afirmó que puede caerse de nuevo en un formato de pregunta y respuesta. Con respecto la tarea B, se consideró que el juego de rol no es el más apropiado porque condiciona al

¹⁰ Un discurso es coherente en el sentido en que el significado del discurso que se produce no contradiga el conocimiento del mundo que tiene el interlocutor. Es cohesionado en la medida que se utilicen elementos de referencia, haya una progresión temática (Instituto Cervantes, 1996-2014).

candidato en una situación hipotética cuando, en la clasificación, lo que se quiere es conocer al candidato y no que asuma otro rol. De igual forma, se sugirió hacer una revisión de las preguntas ya que las respuestas pueden ser cerradas y cortas. Para la tarea C, se evidenció que el tema (el transporte) no parece ser el apropiado para este nivel, pues un candidato de menor nivel de lengua podría resolver la tarea. En suma, se aconseja subir el nivel de dificultad de esta tarea. Finalmente, se sugiere para todas las tareas, disminuir la cantidad de instrucciones para ganar tiempo y, se sugiere que el entrevistador haga valoraciones al candidato con el fin de motivarlo y promover la actuación.

Por otro lado, durante el pilotaje se observó que se respondía mejor a la tarea B que a la tarea A. Esto se debió a que la tarea B proporcionaba más input para la actuación del candidato. Por lo anterior, se sugirió reevaluar las indicaciones de la tarea A para evitar una producción corta y limitada. Se ve entonces la necesidad de ofrecer un input extra proporcionado por el evaluador o por las indicaciones para fomentar una descripción más amplia por parte del estudiante. Por otra parte, también se observó que la propuesta de prueba oral podría ser poco viable en el contexto del CLAM pues, debido a la cantidad de candidatos que presentaron la prueba de clasificación, la propuesta de prueba oral no clasificaba en el tiempo esperado por el centro y por los candidatos. Es decir, en un rango de 10 a 15 minutos.

Conclusiones

Después de haber hecho un recorrido teórico de la evaluación de la lengua oral y haber realizado un proceso de planificación, diseño, validación y experimentación de una propuesta de prueba oral para la clasificación para CLAM, se puede concluir que la viabilidad es una cualidad preponderante entre las cualidades de una prueba de clasificación por dos razones.

Primera, aunque la literatura presente diferentes tipos y métodos para evaluar la validez, fiabilidad, interactividad, autenticidad e impacto de una prueba, la viabilidad determinará si se cuenta con los recursos humanos (desarrolladores, profesores, expertos y candidatos) y el tiempo para realizar dichos procesos. Es decir, en función de la viabilidad se decidirá qué cualidades se van a evaluar y validar en una prueba.

Segunda, aunque se defina un constructo de lengua comunicativo para evaluar, la viabilidad determinará hasta qué punto se pueden desarrollar tareas evaluativas que den cuenta de esa visión de lengua, pues establecer una conexión entre el constructo a evaluar y los procedimientos para

su evaluación, exige tiempo para el desarrollo y evaluación de las tareas, las especificaciones de la prueba y el entrenamiento de los evaluadores. Puede verse entonces que la viabilidad, como consecuencia de un contexto evaluativo y sus limitaciones, es una cualidad que predomina sobre otras cualidades como la validez, la fiabilidad y la autenticidad.

En lo que concierne las tareas de evaluación en un enfoque comunicativo *fuerte* (enseñanza mediante tareas), se concluye que el enfoque *débil* podría ser la mejor alternativa para evaluar la lengua oral, en una prueba de clasificación, mediante tareas comunicativas por dos razones. En primer lugar, una tarea comunicativa de corte *débil* lleva al candidato a utilizar las habilidades lingüísticas que se quieren valorar para su clasificación. Es decir, el objetivo de la clasificación no es valorar lo que el candidato podría hacer con la lengua en el mundo real (como en un examen de proficiencia y evaluación de dominio), sino valorar las habilidades que se enmarcan en un determinado programa para clasificarlo en un curso apropiado.

En segundo lugar, la aplicación de un enfoque *duro* implica llevar al estudiante a que realice tareas del mundo real en un proceso de interacción y negociación de mensajes. Por consiguiente, se presentan dos dificultades. La primera, la tarea debería darle al candidato un input que le permita facilitar las oportunidades de interacción y negociación con el evaluador. Sobre todo, cuando este último ejerce un poder de control discursivo durante la evaluación, lo que hace que difícilmente el candidato negocie con el evaluador no solo para mostrar sus habilidades con la lengua, sino también para realizar la tarea en sí misma. La segunda, como consecuencia de la viabilidad, la tarea debería clasificar al candidato en todos los niveles del programa en un tiempo máximo de 15 minutos. Por lo anterior, se concluye que una prueba oral de clasificación, en oposición a la literatura, no necesariamente debería estar relacionada con el programa de enseñanza.

Finalmente, con respecto a los criterios de evaluación, se evidencia que la literatura aboga por el uso de criterios de evaluación positivos, lógicos y sencillos. Sin embargo, después de analizar los criterios y los descriptores del MCER y de los exámenes DELE, se propone que para una prueba de clasificación, los criterios deberían ser específicos y detallados para determinar minuciosamente las habilidades del candidato en cada uno de los niveles del programa, pues, como ya se mencionó, el objetivo es clasificar y valorar el dominio lingüístico. Para este nivel de detalle, se aconseja el uso del PCIC y los objetivos comunicativos de los programas de la institución.

REFERENCIAS

- Alderson, J., Clapham, C., y Wall, D. (1995). *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press.
- Bachman, L. y Palmer, A. (2010). *Language Assessment in Practice*. Oxford: Oxford University Press.
- Bachman, L. (1990). *Fundamental Considerations in language testing*. Oxford: Oxford University Press.
- Bachman, L. y Palmer, A. (1996). *Language Testing in Practice: Designing and Developing Useful Language Test*. Oxford: Oxford University Press.
- Bordon, T. (2004). La evaluación de la expresión oral y de la comprensión auditiva. J. Sánchez., y I. Santos. (Ed.), *Vademécum para la formación de profesores* (pp. 983-1003). Madrid, España: Sgel.
- Brown, J. (2005). *Testing in language programs. A comprehensive guide to English language assessment*. New York: McGraw-Hill ESL / ELT.
- Consejo de Europa. (2002). *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Madrid, España: Ministerio de Educación, Cultura y Deporte; Instituto Cervantes. [Título original: *Common European Framework of Reference for Languages Learning, teaching, assessment* (2001), Strasbourg: Council of Europe].
- Estaire, S. (2010). Principios básicos y aplicación del aprendizaje mediante tareas. *Revista e didáctica español como lengua extranjera*. Marco ELE.
- Estaire, S. y Zanón, J. (1990). El diseño de unidades didácticas mediante tareas: principios y desarrollo. *Comunicación, Lenguaje y Educación*, 7.
- Font, J. (1998): “Métodos y enfoques en la enseñanza de lenguas extranjeras”, A. Mendoza (coord.) *Conceptos clave en didáctica de la lengua y la literatura*. Barcelona: Horsori.
- Hymes, D. (1967) Models of the interaction of language and social settings. *Journal of social issues*, 23(2): 8-28.
- Instituto Cervantes. (1997-2016). *Competencia comunicativa, corrección, pronunciación*. Centro virtual Cervantes [versión electrónica]. Madrid, España. Recuperado de http://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/
- Luoma, S. (2004). *Assessing speaking*. Cambridge, United Kindom: Cambridge University Press.
- McNamara, T. (2000). *Language Testing*. Oxford: Oxford University Press.
- Martínez, A (2011). Diseñar la evaluación para el alumnado inmigrante en la ESO: evaluación académica / evaluación de habilidades. Ballano, I (Ed.), *Evaluar la competencia lingüística del alumno inmigrante* (pp. 51-70). Bilbao, España. Universidad de Deusto.
- Martínez, A. (2008). *La evaluación oral: una equivalencia entre las guidelines de ACTFL y algunas escalas del MCER* (Tesis doctoral). Universidad de granada, Granada, España.
- Melero, P. (2008). De los programas nocio-funcional a la enseñanza comunicativa. J. Sánchez Lobato e I. Santos Gallardo (directores), *Vademecum para la formación de profesores*. Madrid. pp. 689-714.

- Mingarro, P. (2008). Las escuelas oficiales de idiomas en Cataluña: marco normativo, currículo y certificación. *Monográficos Marco ELE Evaluación*, (7), 70-77.
- Pastor, S (2003). La evolución del proceso de aprendizaje de segundas lenguas. *Perspectivas teóricas y metodológicas: lengua de acogida, educación intercultural y contextos inclusivos*. Madrid, Comunidad de Madrid, Consejería de Educación. Pp. 503-514
- Pastor, C (2007). Sobre el diseño de pruebas orales. *Actas del Programa de Formación para el profesorado de Español como Lengua Extranjera (Múnich, 2006-2007)*[Versión electrónica]. Múnich: Instituto Cervantes de Múnich. Recuperado de http://cvc.cervantes.es/Ensenanza/biblioteca_ele/publicaciones_centros/PDF/munich_2006-2007/08_pastor.pdf
- Soler, O. (2008). Diferentes enfoques para en la evaluación de lenguas mediante tareas. *Monográficos Marco ELE Evaluación*, (7), 111-121.
- Vásquez, G. (2000). *La destreza oral*. Madrid, Edelsa.

Organizan



Apoyan

