"To MCRE or not to MCRE"
Sobre "sentémonos", una plantilla de evaluación más sencilla y flexible

Kris Buyse
(KU Leuven)

Resumen

Tal como señala Buyse (2012), faltan modelos claros y plantillas que permitan evaluar al mismo tiempo y adecuadamente tanto el output comunicativo como la adecuación a la forma. De ahí que en este artículo[1], después de esbozar el punto de partida y el estado de la cuestión de nuestra investigación, recorreremos la metodología, las preguntas de investigación y los resultados del análisis de un corpus de plantillas de evaluación. Terminaremos por presentar nuestra propia propuesta de plantilla de evaluación, denominada "Sentémonos".

1. Puntos de partida y estado de la cuestión

Evaluar la competencia oral hoy día no es nada fácil (Luoma, 2004; Van Iseghem, 2010), entre otras razones porque los objetivos comunicativos de los cursos de lengua actuales hacen que se pida a los evaluadores que (ya) no solo se evalúe el nivel de lengua (con respecto a los componentes gramaticales, léxicos, fonéticos, etc.) sino *al mismo tiempo* la competencia comunicativa bajo forma de una evaluación del desempeño ("performance assessment") a través de actividades realistas (Parrondo Rodríguez, 2004).

Como los criterios, contenidos y plantillas de la evaluación dependen de los objetivos del proceso de aprendizaje, la importancia respectiva de la competencia comunicativa y del nivel de lengua entre los objetivos determinará también la importancia relativa de ambos aspectos en la evaluación. En un programa de enseñanza superior tal como el de Traductores e Intérpretes en Flandes, por ejemplo, se exige un nivel mucho más alto de corrección y adecuación a la forma que en la enseñanza para adultos. Esto explica por qué, según las preocupaciones expresadas por los profesores belgas en sus contactos con nuestro programa de formación docente, las plantillas de evaluación actuales no parecen adaptarse de manera flexible a las necesidades de cada uno de los cursos de idiomas. De nuestro corpus de plantillas de evaluación[2] se desprende, por ejemplo, que las utilizadas por los profesores de lengua en la educación superior la importancia concedida a la adecuación a la forma parece demasiado alta en comparación con los objetivos de los cursos, mientras que en la educación de adultos está casi ausente.

1. Este artículo es una versión aumentada de Buyse (2012), que se limitaba al contexto del Español para Fines Específicos (EFE). Informaciones más detalladas en inglés también se pueden leer en Buyse (2013, i.p.).
2. Véase el apartado 3.

En el contexto de enseñanza actual, la evaluación no constituye un componente *después de la enseñanza*, sino que forma un *componente propio* de la enseñanza, y uno de los más orientadores —es decir, que permite orientar el aprendizaje del alumno—, a condición de que (i) todos los implicados (profesores y alumnos) dispongan de informaciones detalladas sobre el progreso realizado y los problemas eventuales, y que (ii) el profesor no sólo dirija la transferencia de conocimientos sino que también acompañe al alumno en su búsqueda de informaciones y su construcción de *conocimientos (lingüísticos) y destrezas (comunicativas)* (Sluismans & Dochy, 1998). Por lo tanto, la evaluación no sólo tiene como objeto los conocimientos, sino también las destrezas y actitudes (Keeves, 1994). A través de la evaluación del desempeño ("performance assessment"), se evalúa en qué medida el alumno es capaz de aplicar los conocimientos adquiridos a la hora de solucionar nuevos problemas relacionados con los conocimientos antedichos (Meyer, 1992). Para ello, se le deben poner ejercicios realistas o simulados (Gipps, 1994; Parrondo Rodríguez, 2004).

La integración de la enseñanza en el proceso de enseñanza y aprendizaje ha hecho que se emplee en la literatura inglesa cada vez más el término de "assessment" (seguimiento/evaluación continua) para este tipo de evaluación, porque pone énfasis en la relevancia de un contexto de evaluación auténtico —es decir, que simula la(s) realidad(es) en la que harán falta las los conocimientos, destrezas y actitudes evaluados—, en la importancia de dar comentarios a los aprendientes, en las competencias, en el uso —pero no exclusivo— de simulaciones, en la calidad del input para el aprendiz, etc. (Dochy & Gijbels, 2010).

No obstante, según Gijselaers (2007) las generaciones actuales necesitan *estímulos y recompensas* ("triggers and rewards") para conseguir que los alumnos realicen las actividades propuestas con el objetivo de cumplir con los objetivos de aprendizaje, lo cual se consigue entre otras cosas gracias a la evaluación formativa, que, además de evaluar, constituye una recompensa por el trabajo realizado hasta el momento y al mismo tiempo un estímulo para seguir en la misma dirección o adaptarla. Una evaluación *formativa*[3] se considera hoy día esencial, no sólo para el alumno porque le permite detectar elementos positivos y problemas, además de formular objetivos y caminos concretos para mejorarlos, sino también para el profesor, quien puede/debe guiar al alumno en la medida de lo posible, tomando en cuenta el nivel y el progreso del grupo, así como el del individuo (Van Iseghem, 2010). Según Dochy & Gijbels (2010), el empleo frecuente de la evaluación formativa permite aumentar, ampliar y dirigir la motivación del alumno, además de profundizar el aprendizaje, mientras que una evaluación únicamente sumativa conlleva el riesgo denominado "teaching/learning to the test", es decir: que el objetivo primordial del alumno es aprobar el test/examen y el del profesor dar sus clases en función del test/examen.

3. "La *evaluación formativa* es un proceso continuo de acopio de información sobre el alcance del aprendizaje, así como sobre los puntos fuertes y débiles, que el profesor puede introducir en su planificación del curso y en la retroalimentación efectiva que da a los alumnos. La evaluación formativa se utiliza a menudo en un sentido amplio para incluir información no cuantificable proveniente de cuestionarios y consultas. La *evaluación sumativa* resume el aprovechamiento al final del curso con una calificación. No es necesariamente una evaluación de dominio lingüístico; de hecho, gran parte de la evaluación sumativa es una evaluación referida a la norma, realizada en un momento concreto y de aprovechamiento" (Instituto Cervantes, 2002: 186).

Las plantillas que se usan para evaluar en qué medida el alumno ha alcanzado los objetivos son cada vez más *analíticas*, es decir: que la puntuación final constituye la suma de las notas de varios criterios, como los que enumera el Marco Común de Referencia Europea (MCRE): la Competencia lingüística general, la Riqueza del vocabulario, la Corrección gramatical, el Dominio del vocabulario, el Dominio de la pronunciación... Por su parte, la evaluación *global o sintética*[4] se considera más motivadora, porque permite darle al alumno unos comentarios y observaciones detallados en el marco de una evaluación formativa. La global, por su parte, conlleva un mayor riesgo de que se produzca el "efecto halo", es decir: que la evaluación positiva de un solo criterio influya (demasiado) positivamente la evaluación global (Dochy & Gijbels, 2010).

Otra característica del contexto de enseñanza actual es el alto grado de *transparencia, validez y fiabilidad*[5] que se les impone a los objetivos y la evaluación del curso, y el riesgo de que *se recurran* las decisiones del (comité) evaluador en el caso de que no se perciban como suficientemente transparentes, válidas y/o fiables. Debido a eso, una plantilla de evaluación se elabora preferentemente dentro de un *grupo didáctico*, va acompañada de *una clave* (es decir, una nota explicativa sobre cómo manejar la plantilla) y se evalúa regularmente con el objetivo de mejorarla constantemente (Van Iseghem, 2010; Cassany, 2001). Otra consecuencia práctica es que se considera necesario limitar el número de criterios (en la literatura se suele hablar de un máximo de 5 hasta 7 criterios), para que sea *fácil de usar* o *viable*[6], un cuarto criterio general que aparece cada vez más en la literatura, además de la "santa trinidad" de la transparencia, la validez y la fiabilidad.

Finalmente, la gran mayoría de las plantillas de nuestro corpus se refieren a los niveles de lengua establecidos por el MCRE (Council of Europe, 2001, 2009, 2011). Respondiendo a la invitación del Consejo de Europa, varias agencias nacionales para el desarrollo curricular como SLO en los Países Bajos han creado plantillas de evaluación. En base a las demandas de la actual enseñanza y a los deseos expresados durante nuestras sesiones de formación de profesores, podemos afirmar que los profesores belgas buscan plantillas que sean válidas, fiables y transparentes, además de que reflejan las características y exigencias que acabamos de describir. Por eso, las plantillas que presentamos aquí *se basan en* y *refieren a* las consignas y descriptores del MCRE, pero van más allá, porque aunque con este marco se puede evaluar el output comunicativo y el nivel de lengua a la vez, el margen de flexibilidad es mínimo: cada cambio en los objetivos, el nivel o el tipo de enseñanza requiere que se elabore otra plantilla, de modo que es difícil constatar un denominador común entre las plantillas[7]. Así, la agencia neerlandesa Cito

4. "La *evaluación global* consiste en realizar una valoración sintética tomada en conjunto. En este caso, el examinador pondera distintos aspectos de forma intuitiva" (Instituto Cervantes, 2002: 181).
5. "La validez es el concepto del que se ocupa el Marco. Se puede decir que una prueba o un procedimiento de evaluación tiene validez en la medida en que pueda demostrarse que lo que se evalúa realmente (el constructo) es lo que, en el contexto en cuestión, se debería evaluar y que la información obtenida es una representación exacta del dominio lingüístico que poseen los alumnos o candidatos que realizan el examen. La fiabilidad, por otro lado, es un término técnico. Es básicamente el grado en que se repite el mismo orden de los candidatos en cuanto a las calificaciones obtenidas en dos convocatorias distintas (reales o simuladas) de la misma prueba de evaluación" (Instituto Cervantes, 2002: 177).
6. "Un procedimiento de evaluación también tiene que ser práctico, *viable*; la viabilidad tiene que ver concretamente con la evaluación de la actuación. Los examinadores trabajan bajo la presión del tiempo; sólo ven una muestra limitada de la actuación y existen límites definidos del tipo y del número de categorías que pueden manejar como criterios" (Instituto Cervantes, 2002: 178)
7. Véanse por ejemplo las plantillas en Cito/Slo (2011: 97 et ss.).

señala que, como el MCRE distingue entre la expresión, la interacción y la comprensión orales, cada una requiere también su propia evaluación (Cito/Slo 2011: 23). Se ofrece la posibilidad de una evaluación integrada de las destrezas, pero en tal caso "se debe entender que en el caso de un mal desempeño de los estudiantes es difícil averiguar lo que puede ser la causa" (ibídem), lo cual reduce seriamente las posibilidades de feedback en unas plantillas que por defecto no prestan mucha atención a la combinación de la evaluación formativa y sumativa, ni al binomio analítico/sintético. Además, según los mismos autores, la evaluación por medio del MCRE constituye una tarea compleja, porque existen 13 criterios entre los que hay que elegir. Por otro lado,

> ¿Qué pasa con una prueba en la que el profesor quiera evaluar destrezas y los componentes lingüísticos al mismo tiempo? ¿Puede el profesor pretender que su evaluación sea coherente con la visión del MCRE? Pues es importante señalar que el énfasis está en la evaluación de destrezas como leer, escuchar, hablar, conversar y escribir. Si esto no es el caso, entonces se hace difícil sostener que la evaluación se haya basado en el MCRE (Cito/Slo, 2011: 23, traducción k.b.).

Resumiendo, constatamos que faltan plantillas de evaluación que, además de ser suficientemente válidas, fiables, transparentes y fáciles de usar, reflejen adecuadamente las características y exigencias del marco de enseñanza que acabamos de describir, es decir: que (1) evalúen al mismo tiempo el output comunicativo y el nivel de lengua, (2) combinen las ventajas de las evaluaciones analítica y sintética, (3) tengan suficiente flexibilidad para adaptarse a cualquier tipo de enseñanza de lenguas y sus objetivos correspondientes, (4) permitan combinar una evaluación sumativa y formativa de conocimientos, destrezas y actitudes, para (5) convertirse así en un instrumento fiable y suficientemente detallado para dar feedback al alumno.

2. Metodología y preguntas de investigación

Con el objetivo de comprobar la eficacia y adecuación de las plantillas de evaluación actuales, constituimos entre 2009 y 2011 un corpus de 100 plantillas de evaluación provenientes de todos los tipos de enseñanza, niveles y lenguas en Bélgica, tanto de cursos de ELE general como de EFE[8].

Partiendo de nuestro estado de la cuestión, formulamos las siguientes preguntas de investigación:

i) ¿Son suficientemente válidas las evaluaciones? Es decir, ¿comprueban que se hayan cumplido los objetivos?
ii) ¿Son fiables? Para que así sea, un alumno tendría que obtener aproximadamente la misma puntuación, cualquiera que sea el colega que puntúe.
iii) ¿Son transparentes (se puede comprobar en detalle cómo se ha llegado a tal puntuación)?

8. Tipos de enseñanza: enseñanza de adultos 45%, enseñanza superior 35% y enseñanza secundaria 20%. Niveles: A 30%, B 50%, C 20%. Lenguas: español 40%, francés 30%, inglés 15%, neerlandés 10%, otras 5%. ELE general 70%, EFE 30%.

iv) ¿Toman en cuenta el progreso realizado por el/la aprendiente? ¿Son formativas y/o sumativas?
v) ¿Son fáciles de usar o "viables" las plantillas empleadas?
vi) ¿Cómo se relacionan con el MCRE?

En la etapa siguiente de la investigación, analizamos el corpus a partir de las preguntas de investigación (para los resultados, véase el apartado 3).

Finalmente, en base a los resultados de nuestro análisis y siguiendo la invitación de los autores del *Manual for relating language examinations to the Common European Framework of Reference for Languages* (CEFR), cuyo "primary aim ... is to help the providers of examinations to develop, apply and report transparent, practical procedures in a cumulative process of continuing improvement in order to situate their examination(s) in relation to the CEFR" (Council of Europe, 2009: 1), hemos desarrollado una nueva plantilla, denominada "Sentémonos", que fue presentada en diversos foros nacionales e internacionales, los cuales se aprovecharon para hacer encuestas entre los participantes con el objetivo de conocer sus opiniones con respecto a las preguntas que acabamos de listar.

3. Análisis del corpus

El análisis del corpus nos permitió sacar las siguientes conclusiones.

i) En el 97% de los casos las plantillas son más bien *analíticas* (véase el apartado 2), por lo que es más alto el grado de transparencia y fiabilidad que en el caso de una evaluación puramente global o sintética, que por su parte tiene la ventaja de ser más fácil de usar, porque es sencilla y rápida. Sin embargo, la presencia e importancia de los criterios manejados no siempre corresponden con los objetivos del curso, de modo que la validez es más bien baja. Por ejemplo, muchas veces uno de los objetivos importantes del curso es alcanzar una pronunciación poco marcada, pero esto no se refleja en los criterios de la evaluación, donde ésta no se evalúa o no con la misma preponderancia que la que uno esperaría a partir de los objetivos; otras deficiencias frecuentes, además de la imprecisión hasta la falta de objetivos explícitos del curso, son la falta de correlación entre los objetivos comunicativos del curso y una evaluación mayoritariamente basada en la corrección lingüística.

   Finalmente, con respecto a la corrección lingüística, se constata que en la mitad de los casos sólo se evalúa la adecuación a la forma, sin controlar la variación, riqueza y precisión, tanto de las estructuras léxicas y gramaticales como de las pragmático-discursivas; con respecto a las destrezas, en menos de la mitad de los casos se evalúa, además de la expresión oral, también la *comprensión* auditiva, a pesar de su presencia general entre los objetivos de los cursos.

ii) Sólo un 10% va acompañado de una clave, de modo que es muy difícilmente manejable por otro evaluador. De hecho, el 87% de las plantillas no se habían elaborado en equipo, de manera que cada profesor individual manejaba su(s)

propia(s) plantilla(s), lo cual resulta negativo para el grado de *fiabilidad* de la evaluación.

iii) Gracias a su carácter analítico, la gran mayoría de las plantillas (97%) son *relativamente transparentes*. El carácter *relativo* se debe al grado de transparencia y fiabilidad de las escalas utilizadas: calificaciones generales como "bien", "muy bien", etc. o "3", "4", etc. sobre una escala de 5 no son unívocas (cada evaluador las interpretará a su manera), de modo que requieren en realidad una nota explicativa precisa, o se sustituirían mejor por escalas específicas por componente y nivel (véase el ejemplo 2 para una criteriología de la fluidez en el nivel A, elaboración propia):

> Ejemplo 1. Criteriología para la fluidez en el nivel A.
> **2.** El alumno lo entiende todo. No le cuesta mucho conversar y toma iniciativas para seguir conversando. **1,5-1.** El alumno no lo entiende todo, pero formula correctamente las preguntas necesarias para entenderlo finalmente. Le cuesta conversar (muchas pausas, reformulaciones...), pero toma iniciativas para seguir conversando. **0.** El alumno lo entiendo más o menos pero no pide más explicaciones. No consigue conversar y tampoco toma muchas iniciativas para seguir conversando.

En una minoría de casos (2%), se calcula el número de errores por criterio, una intervención a primera vista positiva para la transparencia; sin embargo, no queda claro si se cuenta el número de errores o más bien el número de *tipos de errores* (puesto que no se debería sancionar más de una vez el mismo error), y tampoco es transparente la relación entre la sanción basada en el número de problemas y la cantidad del output comunicativo del alumno (uno que habla poco tendrá menos "oportunidades" para cometer errores).

iv) En una gran mayoría de las plantillas no queda claro en qué medida se usan sólo para fines sumativos o también tienen un carácter *formativo*: son escasas (8%) las que incluyen referencias a (y bonificaciones para) el progreso, el comportamiento y la actitud del alumno a lo largo del período formativo como su participación y colaboración dentro y fuera de la clase, la realización de las tareas, su disponibilidad para aprender y mejorar, etc.; además, muy pocas plantillas se usan para darles comentarios detallados a los aprendientes después del examen o prueba. Casi no hay referencias al uso de un portafolio que dé lugar a una recompensa en la medida en la que los problemas registrados por el alumno se hayan resuelto (sistema de bonificación; la literatura está a favor del portfolio a condición de que el modelo y el sistema estén claros, véase Dochy & Gijbels, 2010).

v) *La facilidad de uso* es muy variable, pero se constata que es inversamente proporcional a la fiabilidad, la validez, el carácter formativo y la transparencia: éstos suben a medida que se añaden más criterios, subcriterios y escalas, pero baja la facilidad de uso, aunque el punto de inflexión depende de cada evaluador individual.

vi) Finalmente, una referencia a los niveles del MCRE está presente en más del 80% de las plantillas, pero sólo en el 2% de ellos vemos la integración de otros aspectos del MCER. Los participantes en la encuesta opinan que el Marco es demasiado complejo y rígido para generar una sola plantilla flexible que combine la adecuación la forma y la adecuación comunicativa, además de integrar conceptos clave de la actual práctica de evaluación como la evaluación formativa/sumativa y analítica/sintética.

Aunque existen otras preocupaciones con respecto a los niveles y los descriptores del MCRE, es evidente que éste constituye un importante paso en la búsqueda global de un conjunto de herramientas de evaluación válida, fiable y transparente. Según el Council of Europe (2009: 3),

> The aim of the CEFR is to facilitate reflection, communication and networking in language education. (...) Thus there is no need for there to be a conflict between on the one hand a common framework desirable to organise education and facilitate such comparisons, and on the other hand the local strategies and decisions necessary to facilitate successful learning and set appropriate examinations in any given context.

4. "Sentémonos": propuesta de plantillas flexibles

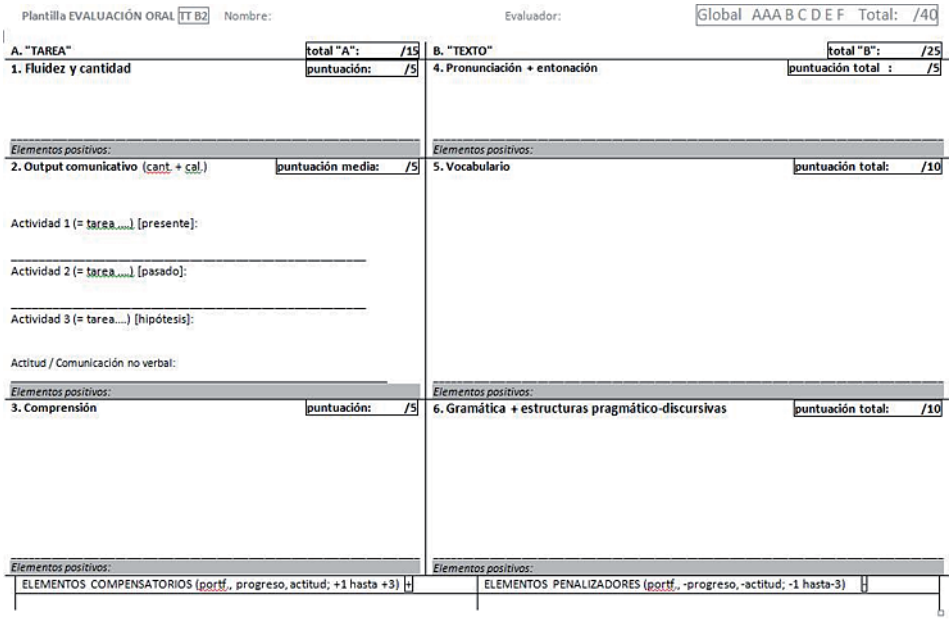
A la luz de los problemas descritos en el apartado anterior, elaboramos unas plantillas flexibles, es decir: adaptables a todo tipo (para fines generales o específicos) y nivel de curso en función de sus objetivos respectivos. Otro objetivo crucial de nuestras plantillas, denominadas "Sentémonos", es que se ofrezca una solución para el desafío descrito en la introducción, a saber: evaluar *al mismo tiempo* la competencia comunicativa y el nivel de lengua a partir de la importancia asignada a cada uno de estos criterios en los objetivos del curso, una *misión casi imposible* según Van Iseghem (2010) y que quedaba por realizar según nuestro análisis del corpus de plantillas. La solución que se propone aquí es una división en dos columnas, una para evaluar los componentes formales (o "Texto": ¿qué sabe el alumno?) y otra para el output comunicativo (o "Tarea": ¿qué hace el alumno?), respectivamente (véase la Tabla 1). Los criterios dentro de cada columna son adaptables a los objetivos de cada curso, aunque parece indispensable que además de la expresión se evalúe también la comprensión. El criterio del "output comunicativo" evalúa el grado de adecuación comunicativa de la expresión e interacción oral (el "saber hacer"), mientras que los criterios del vocabulario, la gramática y la pragmática evalúan en qué medida se recurre para ello a las formas lingüísticas adecuadas (el "saber"), es decir: el léxico adecuado, las formas gramaticales convenientes, además del registro, la coherencia y la cohesión, las estructuras de compensación e interacción, etc. De esta manera se reduce el número de criterios a seis (dos veces tres) y no hace falta elaborar una plantilla aparte para la expresión oral y la interacción oral, respectivamente. El peso relativo de cada columna y de cada criterio depende de los objetivos del curso y del nivel: el del "texto" podría ser mayor en una carrera de filología o de traducción e interpretación que en un curso de EFE, por ejemplo, puesto que en las dos primeras el objetivo es formar a especialistas de la lengua.

Tabla 1. "Tarea" y "texto"

tarea		Texto	
output comunicativo		pronunciación	
fluidez		vocabulario	
comprensión		gramática & pragmática	
	total		total
	%		%
total general			

Otro aspecto fundamental es el aspecto formativo y motivador, incluyendo por ejemplo elementos positivos por criterio y abajo una referencia al proceso de aprendizaje del alumno, con un sistema de bonificación, acompañado o no por otro de penalización por una actitud y progreso negativos (véase la Tabla 2, una plantilla para las competencias orales en el curso de Traducción e Interpretación, nivel B2).

Tabla 2. "Sentémonos": plantilla para las competencias orales en el curso de Traducción e Interpretación, nivel B2.

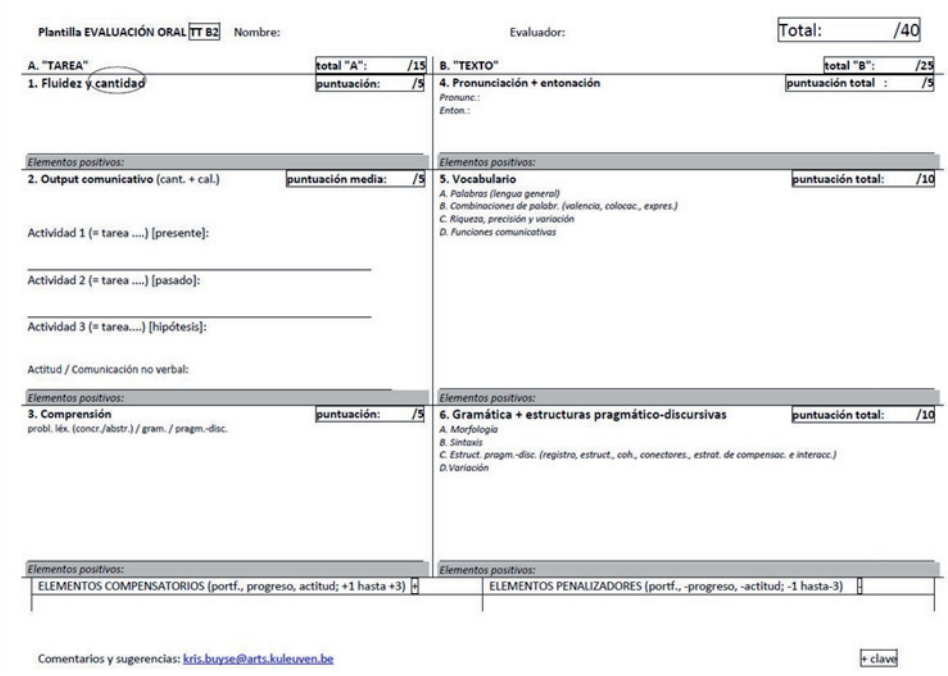


En esta tabla el evaluador tiene la posibilidad de evaluar en dos fases (arriba a la derecha en la Tabla 2): primero apunta durante la prueba oral los elementos positivos y negativos por categoría, sin puntuarlos, pero dando una nota global, intuitiva; en una segunda fase (por ejemplo, durante el tiempo que recibe el alumno siguiente para escuchar y/o ver un fragmento o para leer instrucciones y preparar la respuesta, o bien más tarde cuando haya más tiempo) el evaluador puntúa las categorías según el sistema analítico detallado en una clave. En el caso de que haya una discrepancia importante

entre la evaluación global y la analítica, hace falta reconsiderar el caso y buscar las posibles causas de la diferencia. Del feedback que estamos recibiendo al final de nuestros talleres y de evaluadores que están usando estas plantillas se desprende que este procedimiento por fases genera más confianza en la fiabilidad de la evaluación y que es fácil de usar, porque durante una prueba oral es difícil apuntar observaciones y puntuar adecuadamente al mismo tiempo; así, la puntuación se realiza de manera más fiable y el riesgo de una discrepancia entre la evaluación global y la analítica es menor, lo que se confirma en la realidad, según los testimonios recibidos.

En función de las preferencias del evaluador, es posible añadir más "capas de evaluación": en la Tabla 3 se han añadido más subcriterios a las categorías (para más capas, véanse Buyse 2012, 2013).

Tabla 3. "Sentémonos": plantilla para las competencias orales en el curso de Traducción e Interpretación, nivel B2. Segunda "capa de evaluación".



Plantilla EVALUACIÓN ORAL TT B2 Nombre: Evaluador: Total: /40

A. "TAREA" total "A": /15
1. Fluidez y cantidad puntuación: /5

B. "TEXTO" total "B": /25
4. Pronunciación + entonación puntuación total : /5
Pronunc.:
Enton.:

Elementos positivos:
2. Output comunicativo (cant. + cal.) puntuación media: /5

Elementos positivos:
5. Vocabulario puntuación total: /10
A. Palabras (lengua general)
B. Combinaciones de palabr. (valencia, colocac., expres.)
C. Riqueza, precisión y variación
D. Funciones comunicativas

Actividad 1 (= tarea ....) [presente]:
Actividad 2 (= tarea ....) [pasado]:
Actividad 3 (= tarea....) [hipótesis]:
Actitud / Comunicación no verbal:

Elementos positivos:
3. Comprensión puntuación: /5
probl. léx. (concr./abstr.) / gram. / pragm.-disc.

Elementos positivos:
6. Gramática + estructuras pragmático-discursivas puntuación total: /10
A. Morfología
B. Sintaxis
C. Estruct. pragm.-disc. (registro, estruct., coh., conectores., estrat. de compensac. e interacc.)
D.Variación

Elementos positivos:
ELEMENTOS COMPENSATORIOS (portf., progreso, actitud; +1 hasta +3) +

Elementos positivos:
ELEMENTOS PENALIZADORES (portf., -progreso, -actitud; -1 hasta-3)

Comentarios y sugerencias: kris.buyse@arts.kuleuven.be

+ clave

A pesar de que sean adaptables al tipo de curso, su nivel, la visión del evaluador y del instituto, es menester actualizar regularmente las plantillas, tomando en cuenta los problemas surgidos anteriormente y las discusiones entre colegas. Una herramienta que permite a cualquier evaluador comprobar en qué medida su evaluación corresponde a la de otros colegas, es WebCef[9], donde uno puede comparar su propia evaluación de una prueba grabada con la de otros colegas.

9. www.webcef.eu

Finalmente, para aumentar la fiabilidad, es esencial que cada versión de una plantilla vaya acompañada de una clave que contenga los objetivos y la materia del curso, el sistema de puntuación, los descriptores, las series de actividades y el sistema de selección, las listas de problemas típicos por nivel.

5. Conclusiones y discusión

En conclusión, el análisis de un corpus de plantillas de evaluación de las competencias orales sugiere que hacen falta, tanto para los cursos de ELE general como para los de EFE, unas plantillas más fiables, válidas y formativas, y que también son mejorables la transparencia y la facilidad de uso. En el contexto de un curso de EFE son problemáticas especialmente las relaciones entre los objetivos comunicativos y su evaluación, entre los conocimientos y las destrezas, y entre la corrección lingüística y el output comunicativo.

Las plantillas flexibles que se han presentado en el apartado anterior se proponen cumplir con estos requisitos, al mismo tiempo que se pueden usar en cualquier curso y nivel, adaptándose a las exigencias del curso y las preferencias del evaluador. Asimismo quieren servir como instrumento formativo con el objetivo de dar la mayor cantidad posible de comentarios a los alumnos sobre su actuación durante el examen oral. En este sentido es más bien un instrumento de "assessment" (véase el apartado 2) que de "evaluación". Aquél tiene su origen en la palabra latina "asidere", lo que significa "estar sentado al lado de alguien" y pone de relieve el proceso de aprendizaje (Dochy & Gijbels 2010). Por lo tanto, la metáfora que los profesores podemos tener en mente a la hora de evaluar, es que más vale sentarnos *al lado del* alumno que *enfrente de* él. Y de ahí el título de "Sentémonos"...

Bibliografía

BUYSE, K. (2012): «Criterios y plantillas para la evaluación del español para fines específicos», en *El español de las profesiones. Artículos seleccionados del IV Congreso Internacional de Español para Fines Específicos. Congreso Internacional de Español para Fines Específicos*, Amsterdam, 18-19 November 2011, Madrid: Ministerio de Educación, Cultura y Deporte, 186-200

BUYSE, K. (2013, i.p.): «"Come and sit here next to me". Towards a communicative assessment of oral language skills», *Language Learning in Higher Education, 2*.

CASSANY, D., M. LUNA y G. SANZ (2001): *Enseñar lengua*, Barcelona: Grao.

CITO/SLO (2010): *Toetsen en beoordelen met het ERK*, Arnhem/Enschede.

COUNCIL OF EUROPE (2001): *Common European Framework of Reference for Languages: Learning, Teaching, Assessment,* Cambridge: Cambridge University Press.

COUNCIL OF EUROPE (2009). *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR): A Manual,* Strasbourg: Council of Europe.

COUNCIL OF EUROPE (2011): *Manual for Language Test Development and Examining*, Strasbourg: Council of Europe.

DOCHY, F., y D. GIJBELS (2010): «Evaluatie», en S. Janssens (dir.), *Leren en onderwijzen,* Leu-

ven: Acco, 121-175.
GIJSELAERS, W. (2007): «Talking about My Generation», Keynote at EARLI 07: European Practice-based and Practitioner Research conference on Learning and Instruction for the new generation, Maastricht, 14-16 November.
GIPPS, G.V. (1994): *Beyond Testing: Towards a Theory of Educational Assessment,* London: The Falmer Press.
INSTITUTO CERVANTES (2002): *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación,* Madrid: Instituto Cervantes.
KEEVES, J.P. (1994): «Methods of assessment in schools», en T. Husén y T.N. Postlewhaite (dirs.), *International Encyclopedia of Education,* Oxford/New York: Pergamon Press, 362-370.
LUOMA, S. (2004): *Assessing Speaking,* Cambridge: Cambridge University Press.
MEYER, C.A. (1992): «What's the difference between authentic and performance assessment», *Educational Leadership,* 49, 39-40.
PARRONDO RODRÍGUEZ, J.R. (2004): «Modelos, tipos y escalas de evaluación», en J. Sánchez Lobato e I. Santos Gargallo (dirs.), *Vademécum para la formación de profesores. Enseñar español como segunda lengua (L2)/lengua extranjera (LE),* Madrid: SGEL, 967-982.
SLUISMANS, D., y F. DOCHY (1998): «Alternatieve toetsmethoden in studentgericht onderwijs», *Tijdschrift voor Hoger Onderwijs*, 16 (4), 229-233.
VAN ISEGHEM, J. (2010): *Didactiek mondelinge vaardigheden*, Leuven: KU Leuven, curso del Máster de Didáctica de Lenguas.