EL PAPEL DE LA NEGOCIACIÓN DIALÓGICA EN LAS EVALUACIONES ORALES DE LA PRUEBA FINAL DE BACHILLERATO

JESÚS GARCÍA LABORDA
Universidad de Alcalá, Madrid
ANA ISABEL MUÑOZ ALCON
Universidad Católica, Ávila

RESUMEN

Durante años se han seguido dos modelos de evaluación de lenguas extranjeras principales en el mundo. Por un lado, está el modelo socio-cognitivista de Weir (2005) y por otro, el de Chapelle, Enright y Jamieson (2007). Ninguno de los dos considera la co-construcción discursiva ni sus efectos positivos en la actuación del candidato. Este estudio muestra que la negociación dialógica en la Zona de Desarrollo Próximo (Poehner, 2007) favorece la producción oral en exámenes. Además, expone la comparación de resultados entre un grupo control de la Comunidad Andalucía y otro de las Comunidades de Castilla La Mancha y Madrid mediante grabaciones realizadas a 60 estudiantes en entrevistas orales pareadas. Los resultados muestran que las interacciones entre los dos candidatos son más efectivas que los diálogos entrevistador-candidato, lo que podría afectar a la gestión y realización de la Prueba Final de Bachillerato que comenzará en 2017.
Palabras clave: evaluación, exámenes pareados, co-construcción discursiva, ZDP.

ABSTRACT

For years two main approaches to assess foreign languages have been followed in the world: on the one hand, Weir's socio-cognitive model (2005) and the other model, also cognitive, is the Chapelle, Enright, and Jamieson's (2007). Neither considers the discursive co-construction with consequent effects on improving the candidate's performance. It shows as well that the dialogic interchange on the Zone of Proximal Development (Poehner, 2007) favors the oral production in tests. The study compares the results from a control group of the Community of Andalucía and another of Castilla La Mancha and Madrid recordings made by 60 students in paired oral interviews. The results show that interactions between the two candidates are more effective than the dialogues interviewer-candidate and they may actually affect the management and implementation of the Final Test High School (Prueba Final de Bachillerato) beginning in 2017.
Keywords: evaluation, semidetached exams, constructing discourse, ZDP.

1. INTRODUCCIÓN

La LOMCE supone una ruptura con la evaluación de bajo impacto tradicional del sistema educativo español. El cambio venía ya reclamándose desde hace mucho tiempo, especialmente en lo referente a la Prueba de Acceso a la Universidad (PAU). García Laborda (2010, 2012), Amengual Pizarro (2009) y Díez Bedmar (2011), entre otros, han afirmado que esta necesidad es imperiosa para la mejora del aprendizaje de lenguas extranjeras en España. Lo cierto es que, como García Laborda ha señalado, el salto de la PAU a la nueva prueba denominada Prueba Final de Bachillerato (PFB) conllevaría una serie de ventajas entre las que destaca una mayor viabilidad para el desarrollo de las tareas orales dentro de la sección de inglés.

En el año 2012, el Instituto Nacional de Evaluación Educativa (INEE) realizó un estudio piloto para analizar la introducción de una prueba oral en la sección de inglés de la Prueba de Acceso a la Universidad (PAU), que hasta ahora, sólo consta de prueba escrita. En total, 1.033 estudiantes fueron evaluados en un examen de dos partes cuyas tareas consistían en una conversación informal examinador-alumno y la descripción de una imagen. Los resultados publicados por el Ministerio de Educación, Cultura y Deporte (MECD) muestran que más de la mitad (61,3%) de los participantes poseían un nivel B1 en el Marco Común Europeo de Referencia para las Lenguas (MCERL). El informe no encontró diferencias significativas en referencia al sexo o la edad de los estudiantes (Tabla 1).

	Criterio					Total
	Alcance	Corrección gramatical	Fluidez	Interacción	Coherencia	
Parte 1	65.08	60.39	64.66	68.17	68.51	63.65
Parte 2	61.06	55.53	62.81	64.49	65.66	60.80
Total	60.22	54.77	61.89	63.82	64.57	61.39

Tabla 1. Porcentaje de alumnos que alcanzan un nivel B1 según tarea, criterios de evaluación y el total del examen. Fuente: INEE.

Se encuentran, sin embargo, algunas deficiencias en el estudio:

- Los resultados podrían estar mediados por la falta de experiencia en evaluación oral, así como la carencia en formación específica o incluso los juicios previos de los evaluadores –que eran profesores voluntarios de secundaria- y el conocimiento por parte de los estudiantes –también voluntarios, por lo que se puede deducir que su nivel era mayor al de la media- de que los resultados no afectarían a su evaluación final, entre otros diversos factores.
- El INEE no detalló en su informe cómo se midieron y valoraron los criterios de corrección.
- No se publicaron tampoco los resultados diferenciados por territorio –esto último quizás motivado por las diferencias en el reparto de competencias entre las Comunidades Autónomas participantes).

Además, la evaluación individualizada parece no ser un método adecuado por varias razones:

a. Los alumnos españoles raramente tienen evaluaciones orales;
b. El trabajo en el aula cuando hay práctica oral no es individual;
c. El tipo de evaluación del INEE no es realista, pues no se asemeja a las actividades orales que los profesores realizan actualmente ni tiene en cuenta las circunstancias reales de las aulas.

Por todo esto, la validez del constructo de la prueba del INEE está en entredicho, y aplicarla como tal a la nueva PFB sería un error, pues no permitiría realizar las inferencias necesarias acerca de los examinados. Posiblemente se debe a que se ajusta, sin embargo, al modelo socio-cognitivo generalmente utilizado en las Escuelas Oficiales de Idiomas, en TOEFL (Chapelle, Enright, Jamieson's, 2007) o por el Cambridge Board of Examinations (modelo que ha adaptado el British Council para sus exámenes como APTIS) (Weir, 2005). Sin embargo, hay que tener en cuenta que incluso el Cambridge Board of Examinations introdujo recientemente las entrevistas pareadas en sus exámenes, ya que se echaba de menos en estos la interacción entre pares. Son varios los autores que señalan algunas de sus ventajas: por ejemplo, Van Compernolle (2011) afirma que la entrevista oral pareada ayuda a superar ciertos problemas como la ansiedad y McNamara (1996) defiende que los exámenes actuales se fijan mucho en el conocimiento individual con consecuencias negativas.

Efectivamente, los modelos cognitivos resultan competitivos y la comunicación real raramente lo es. Por lo tanto, resulta necesario fijarse en modelos en los que los alumnos interactúen durante el examen. Poehner (2008) sugiere la Evaluación Dinámica, modelo de evaluación basado en la interacción de dos individuos en la Zona de Desarrollo Próximo, siendo la tarea del evaluador mediar activamente modificando y mejorando el input del examinado, hasta que su producción sea comprensible para el alumno. La idea, desafortunadamente, se parece mucho a la función del profesor en la teoría del ínput de Krashen (1985). Sin embargo, la entrevista pareada deja de lado el aspecto de controlador de la entrevista que aparece en el trabajo de Poehner

(2008), y también de Lantolf y Poehner (2011), y permite acceder a un paradigma en el que los actores que interactúan en la Zona de Desarrollo Próximo (ZDP) compartan intereses, mutuo reconocimiento y ayuda. Entre iguales, no existe la superioridad del evaluador, la interacción es más libre y está más acomodada a las necesidades de los estudiantes. Por lo tanto, la mediación en la ZDP es más constructiva (tanto productivamente como de manera interaccional).

2. INVESTIGACIÓN

Con el fin de investigar el efecto de la entrevista pareada y ver si efectivamente era más productiva, se realizó un estudio entre alumno universitarios de primer semestre de las universidades de Jaén (Psicología) y Alcalá (Magisterio). Mientras que los primeros no recibieron instrucciones precisas de interacción, los segundos sí las recibieron.

Con el objetivo de facilitar la evaluación, se evaluó previamente el nivel global de cada alumno, y posteriormente sus entrevistas fueron analizadas teniendo en cuenta diferentes criterios. En este trabajo nos centramos en el criterio de interacción (tabla 2) y de fluidez (tabla 3). Para ello se utilizó un baremo de 0 a tres siendo 0=muy deficiente, 1=no alcanza una puntuación aceptable para su nivel en el MCERL, 2=alcanza una puntuación aceptable para su nivel en el MCERL, 3=obtiene una marca excelente para su nivel en el MCERL.

2.1 Resultados

La tabla 2 muestra los resultados obtenidos en Jaén (Psicología) y en Alcalá (Magisterio).

			Interaction				Total
			0	1	2	3	
Titulación	Psicología U. Jaén	Recuento	1	12	17	8	38
		% dentro de Titulación	2.6%	31.6%	44.7%	21.1%	100.0%

		% dentro de Interaction	50.0%	20.0%	21.8%	27.6%	22.5%
	Magisterio Primaria U. Alcalá	Recuento	0	14	32	6	52
		% dentro de Titulación	.0%	26.9%	61.5%	11.5%	100.0%
		% dentro de Interaction	.0%	23.3%	41.0%	20.7%	30.8%

Tabla 2. Calificaciones obtenidas por los dos grupos en interacción. (Fuente propia)

			Fluency				Total
			0	1	2	3	
Titulación	Psicología U. Jaén	Recuento	1	9	17	11	38
		% dentro de Titulación	2.6%	23.7%	44.7%	28.9%	100.0%
		% dentro de Fluency	100.0%	18.4%	20.2%	33.3%	22.8%
	Magisterio Primaria U. Jaén	Recuento	0	12	31	7	50
		% dentro de Titulación	.0%	24.0%	62.0%	14.0%	100.0%
		% dentro de Fluency	.0%	24.5%	36.9%	21.2%	29.9%

Tabla 3. Calificaciones obtenidas por los dos grupos en Fluidez. (Fuente propia)

Dada la limitación de la muestra, este estudio piloto muestra de manera no concluyente unos resultados ligeramente favorables a favor de la Universidad de Alcalá. Aunque son homogéneos en los extremos, no lo son en los valores centrales, especialmente en lo referente a la fluidez. Las videograbaciones muestran que mientras las actividades en pares de la Universidad de Jaén suelen carecer de mutuo apoyo y cooperación, las de la Universidad de Alcalá son mucho más cooperativas, estableciéndose tres tipos fundamentales de actuaciones positivas: apoyo, aclaración y compleción. El apoyo se produce cuando el alumno no conoce un término, por ejemplo, y se lo ofrece a su interlocutor continuando con la conversación; la aclaración surge cuando el compañero

explicita una idea de su interlocutor que ha quedado confusa; y finalmente la compleción se da cuando colaborativamente se completan frases que individualmente son incapaces de concluir.

En este punto, y a falta de incrementar el estudio, nuestras observaciones muestran que el apoyo se da a todos los niveles, mientras que los otros dos tipos de co-construcciones dialógicas se dan especialmente entre participantes de distintos niveles y son generalmente iniciadas por el alumno con mayor nivel. Sin embargo, como se apuntó anteriormente, nos encontramos en una fase inicial del estudio, y en los próximos meses se elaborarán informes más detallados.

3. DISCUSIÓN Y CONCLUSIONES

Del estudio se puede concluir, primero, que fomentar un clima cooperativo entre los dos candidatos favorece y promueve una comunicación más eficaz y mejores resultados en el test. En la Figura 1 se muestran las relaciones que se establecen en las tareas llevadas a cabo. Evidentemente, las relaciones más fuertes y estrechas no se establecen en la primera tarea entrevistador-estudiante, sino en la segunda llevada a cabo entre los dos estudiantes. Esto también se pudo ver en varias entrevistas realizadas en Penn State en el programa de verano del año 2013, donde se realizaron un total de 20 entrevistas. En varias ocasiones hemos podido comprobar que el alumno muestra dos comportamientos diferenciados en las entrevistas pareadas dependiendo de si el interlocutor es el entrevistador o un compañero. Así, en las entrevistas de Jaén y Alcalá, los estudiantes se muestran diferentes en la relación 1 que en la 2 (Figura 1).

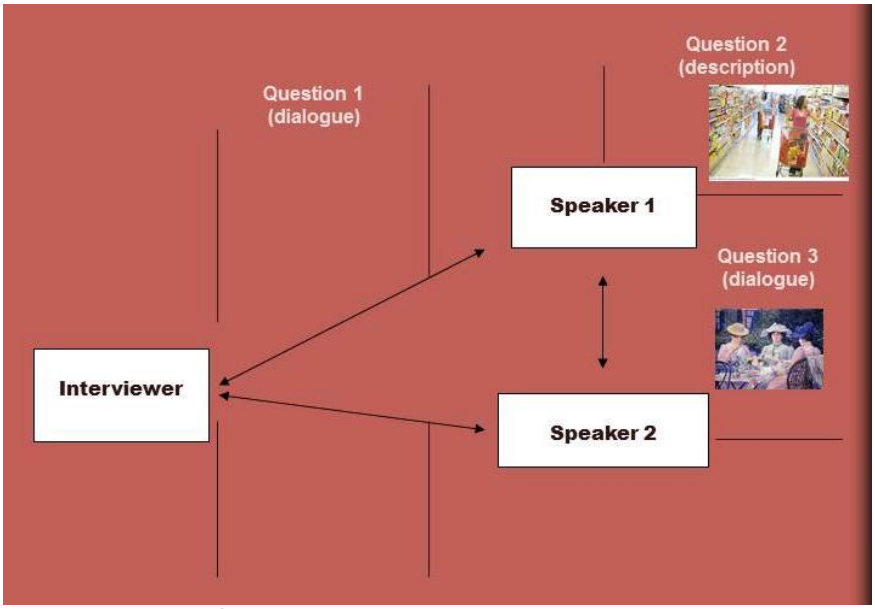

Figura 1. Relación de tareas en entrevistas orales den el proyecto OPENPAU

Ciertamente como cualquier estudio piloto, este trabajo tiene algunas limitaciones. La primera es el tamaño de la muestra, que trataremos de solventar ampliando el número próximamente. La segunda y más evidente, considerar en qué medida afectan realmente los criterios de fluidez y percepción a la nota global propuesta por el evaluador. La tercera y última, la falta de conocimiento sobre los efectos reales que produce el hecho de simultanear la evaluación de dos niveles próximos, pudiendo alterar la percepción del entrevistador. Si bien es cierto que cuando hay dos niveles muy diferenciados el entrevistador es capaz de percibir ambos, podría ser que con niveles más cercanos no se apreciara la diferencia. Aunque este estudio contó con dos evaluadores en todas las entrevistas, siempre hay uno que suele tender a ser el evaluador principal, incluso cuando ambos cuentan con amplia preparación y experiencia. A la vista de estos datos, simplemente cabe preguntarse si las construcciones dialógicas realmente implican una mayor competencia en un contexto con menor ansiedad, o si realmente muestran indicaciones de un constructo comunicativo que podría desarrollarse en el futuro.

En nuestra opinión, estas preguntas son de respuesta obligatoria si este modelo va a ser utilizado en la futura Prueba Final de Bachillerato (PFB), consiguiendo así, finalmente, que los exámenes tengan una influencia real en cómo se enseñan y aprenden las lenguas extranjeras en España.

AGRADECIMIENTOS

Los autores quisieran expresar su agradecimiento al Ministerio de Ciencia e Innovación del Gobierno de España por la subvención recibida, dentro del Plan Nacional de I+D+I en su convocatoria de 2011, para llevar a cabo el proyecto OPENPAU (Orientación, propuestas y enseñanza para la sección de inglés en la Prueba de Acceso a la Universidad, FFI2011-22442).

REFERENCIAS BIBLIOGRÁFICAS

Amengual Pizarro, M. 2009. "Does the English Test in the Spanish University Entrance Examination influence the teaching of English?", *English Studies*, 90 (5): 582-598.

Chapelle, C., Enright, M., & Jamieson, J. (Eds.) 2008. *Building a validity argument for TOEFL*. New York: Routledge/Taylor and Francis Group.

Díez Bedmar, M.B. 2011. "The English exam in the University Entrance Examination: an overwiew of studies", *Revista Canaria de Estudios Ingleses*, 63: 101-112.

García Laborda, J. 2010. "¿Necesitan las universidades españolas una prueba de acceso informatizada? El caso de la definición del constructo y la previsión del efecto en la enseñanza para idiomas extranjeros", *Revista de orientación y Psicopedagogía*, 21(1):71-80.

García Laborda, J. 2012. "De la Selectividad a la Prueba de Acceso a la Universidad: Pasado, presente y un futuro no muy lejano", *Revista de Educación*, 357, 17-28.

Krashen, S.D. 1985. *The Input Hypothesis: Issues and Implications*, New York: Longman

Lantolf, J. P., Poehner, M. E. (2011). *Dynamic Assessment in the Foreign Language Classroom*. A Teacher's Guide. 2nd

Edition, State College: The Pennsylvania State University Press.
McNamara, T. F. 1996. *Measuring second language performance*, London: Longman.
Poehner, M. E. 2008. *Dynamic assessment: A Vygotskian approach to understanding and promoting L2 development*, New York: Springer.
Van Compernolle, R. A. 2011. "Responding to questions and L2 learner interaction competence during language proficiency interviews: A microanalytic study with pedagogical implications", J. K. Hall, J. Hellerman, S. Pekarek Doehler (Eds.), *L2 interactional competence and development*, Bristol: Multilingual Matters.