

Análisis de la variable *tarea* en la cantidad de producción oral de tests en inglés

Marcos Peñate | Universidad de Las Palmas de Gran Canaria

En este artículo estudiamos los tres tipos de tareas más usadas en los diferentes modelos de pruebas orales (entrevista individual a partir de una foto, de una tira cómica y diálogo en parejas). Para hacer este estudio evaluamos a 244 alumnos de segundo de bachillerato pertenecientes a nueve centros escolares y con un nivel A2 de inglés. La variable que hemos querido analizar al contrastar los tres tipos de tareas es la de la cantidad de la producción oral. Se estudió, a través del test ANOVA, las diferencias significativas entre las medias de palabras producidas por alumno. Dicho test nos reveló que el alumnado tiene una producción significativamente menor cuando la tarea se realiza a través de la entrevista individual con foto.

Palabras clave: *clave: inglés como lengua extranjera, evaluación, tareas de producción oral, fluidez, educación secundaria.*

In this paper we study the three types of tasks most frequently used in the different models of oral exams (one-to-one interview based on a photo, one-to-one interview based on a comic strip, and a dialogue in pairs). In order to carry out this study we assessed 244 second-year baccalaureate students at nine different schools with an A2 level of English. When contrasting the three types of tasks, our aim was to analyse the complexity of the oral production. By means of the ANOVA test, we studied the significant differences between the average number of words produced by the students. The test revealed that students produce a significantly lower amount when the test consists of a one-to-one interview based on a photo.

Keywords: *English as a foreign language, assessment, speaking tasks, fluency, secondary education.*

Dans cet article, nous étudions les trois types de tâches les plus utilisées dans les différents modèles d'épreuves orales (entretien individuel à partir d'une photo ou de vignettes humoristiques ou bien de dialogues en paire). Pour mener à bien cette étude, nous avons évalué 244 élèves de Terminale appartenant à neuf établissements scolaires et ayant un niveau A2 en langue anglaise. La variable que nous avons voulu observer en contrastant les trois types de tâches est celle de la quantité de production orale. Nous avons analysé, à travers le test ANOVA, les différences significatives entre les moyennes de mots employés par apprenant. Ledit test a révélé que les élèves ont une production significativement inférieure quand la tâche est réalisée lors d'un entretien individuel avec une photo.

Mots-clés: *anglais comme langue étrangère, évaluation, tâche de production orale, fluidité, enseignement secondaire.*

1. Introducción

La evaluación de una lengua extranjera es un tema que puede ser abordado desde diferentes perspectivas. La perspectiva que el presente artículo pretende abordar se centra en la evaluación de la expresión oral teniendo como eje de la investigación las tareas utilizadas en dicha evaluación.

Dos han sido las líneas de investigación que han prestado especial atención al uso de diferentes tareas a la hora de evaluar la producción oral. La primera de ellas surge de los estudios sobre evaluación de lenguas extranjeras, mientras que la segunda es parte fundamental de la teoría del *output* comprensible. Desde la perspectiva de la evaluación se ha investigado, entre otros aspectos, el formato de entrevista oral individual (examinador versus alumno) en contraste con el formato en parejas, mientras que desde la teoría del *output* comprensible se ha intentado analizar al papel de la producción oral en el proceso de aprendizaje de una lengua extranjera.

Puesto que nuestro estudio incide en modelos de tareas individuales y en parejas, y además analizaremos la producción oral obtenida, dedicaremos esta sección de revisión teórica a estos dos aspectos.

1.1. Pruebas individuales versus pruebas en parejas

Los trabajos de investigación realizados a fines del siglo pasado (Ross y Berwick, 1992; Young y Milanovic, 1992) ya empezaron a señalar una serie de deficiencias relacionadas con el formato individual de entrevista. Paralelamente, en ese mismo período, empiezan a publicarse artículos, como el de Saville y Hargreaves (1999), en los que se enumeran una serie de ventajas del formato en parejas. Ventajas que

apuntan a una atmósfera más relajada al realizar la tarea y una mayor variedad de lengua.

Puesto que en cierta medida la universidad de Cambridge, a través de sus exámenes oficiales, ha sido una de las abanderadas en la implantación del formato en parejas, queremos mencionar dos estudios realizados por dicha institución (Taylor, 2001), en el que se comparan los dos modelos. El primero de los trabajos realizados fue de tipo cuantitativo y en él se comprueba que las tareas en pareja producen una muestra de lengua más rica y variada. El segundo estudio fue de tipo cualitativo y en él se analiza, a través de una lista de observación, el uso que los alumnos hacen de las diferentes funciones comunicativas. Se comprobó que al interactuar en parejas el uso de dichas funciones era más elevado, alcanzándose la cifra de treinta tipos diferentes.

El inicio de este siglo se ha caracterizado por el aumento de estudios, en especial de casos, en los que se han ido analizando diferentes características del formato en parejas. Así se ha podido constatar que, en las situaciones en las que un alumno tiene un mayor nivel lingüístico que su compañero, se produce una mayor producción oral en general (Norton, 2005), aumento que también se constata en el alumno de nivel más bajo (Davis, 2009).

1.2. Fluidez, precisión y complejidad

En 1975, Wagner-Gough y Hatch proponen una nueva función aplicada a la conversación dentro del proceso de enseñanza-aprendizaje. Estos autores argumentan que la interacción constituye la base del desarrollo de la sintaxis, y no es solamente una oportunidad para la práctica de estructuras gramaticales, como se había defendido hasta entonces. Sin embargo, la idea de que el *output* o el uso de la lengua podría constituir

parte del propio mecanismo de aprendizaje no se contempla con rigor hasta 1985, año en que Swain publica un artículo casi fundacional en el que introduce la noción de *output comprensible*. Trabajos posteriores han ido reformulando dicha hipótesis (Swain y Lapkin, 1995, 1998) y determinando sus funciones fundamentales (Arnaiz y Peñate, 2004): potenciar la fluidez, la concienciación de las lagunas lingüísticas, la comprobación de hipótesis y el metalenguaje.

El concepto de *fluidez* se ha prestado a diversas interpretaciones. Así, por ejemplo, Fillmore (1979: 93) la entiende como la habilidad de llenar el tiempo del discurso, de emplear oraciones coherentes y semánticamente densas, con dominio de los recursos semánticos y sintácticos. Para este autor, la fluidez también implica la habilidad de emplear el vocabulario apropiado en diferentes contextos y la habilidad de ser creativo e imaginativo. Sin embargo, esta manera de entender la fluidez, en la que de hecho se incorpora la precisión, se redefine posteriormente, tal como podemos apreciar en la clasificación hecha por Skehan y Foster (1999: 96–97), quienes distinguen entre fluidez, complejidad y precisión. Por *fluidez* se entiende la capacidad de usar la lengua en tiempo real, dando especial énfasis al significado. La *precisión*, por el contrario, es la capacidad de evitar cometer errores, lo que implica un mayor control en el uso de la lengua. Por fin, se define la *complejidad* como la capacidad de usar un lenguaje más sofisticado, aunque incluya estructuras que aún no se dominen, lo que indudablemente muestra una voluntad de arriesgarse a cometer errores. Vemos, por tanto, que el concepto de complejidad coincide con los de concienciación de las lagunas lingüísticas y comprobación de hipótesis, propuestos desde la teoría del *output comprensible*.

2. Desarrollo del trabajo

2.1. Objetivos

Tradicionalmente, los estudios que han analizado la relación entre diferentes tipos de pruebas y la evaluación se han centrado en el análisis de los resultados obtenidos por el alumnado. No obstante, en este trabajo pretendemos profundizar en los efectos lingüísticos que el uso de diferentes tipos de pruebas orales tiene en la producción oral de los alumnos. En concreto queremos desarrollar el siguiente objetivo: comprobar si se dan diferencias significativas en el número de palabras emitidas en los diferentes tipos de pruebas.

2.2. Sujetos de la investigación

Los sujetos de la investigación fueron alumnos de segundo de bachillerato de centros públicos y privados de la isla de Gran Canaria. Se eligieron al azar nueve centros y posteriormente se dividieron estos en tres grupos, procurando que en cada grupo hubiese un centro público urbano, otro rural y un centro privado. Teniendo en cuenta que cada grupo tiene aproximadamente 30 alumnos, la muestra estaría hipotéticamente compuesta por 270 sujetos. En cada centro se aplicó un solo modelo de prueba y la media de alumnos a los que se aplicó en cada centro fue de 27,1, lo que hizo un total de 244 alumnos (cuadro 1, en la página siguiente).

3. Materiales y recursos

Tras algunos estudios previos para garantizar la fiabilidad y el poder de discriminación de las pruebas, se optó por diseñar tres pruebas diferentes. En las pruebas 1 y 2, había solamente dos protagonistas: el aplicador que hacía las preguntas y el alumno que la respondía. La prueba 3, podemos decir que

Cuadro 1. Alumnos por centro y tipo de prueba.

	Aplicadores	Centros	Alumnos		Total
			Previstos	Aplicados	
Prueba 1 (Foto)	Aplicador 1	1	30	30	84
		2	30	30	
		3	30	24	
Prueba 2 (Cómic)	Aplicador 2	4	30	23	70
		5	30	21	
		6	30	26	
Prueba 3 (Dialogo en parejas)	Aplicador 3	7	30	30	90
		8	30	30	
		9	30	30	

constaba de dos partes fundamentales: en la primera, eran dos los alumnos que se cruzaban entre sí cinco preguntas cada uno, mientras que en la segunda, el aplicador les dirigía cinco preguntas a ambos y estos iban contestando a todas las preguntas de manera individual.

La prueba 1 consistía en la presentación a los alumnos de una fotografía que se les entregaba en el momento de la entrevista, a partir de la cual se les pedía que contestaran verbalmente a cinco cuestiones relacionadas directamente con dicha foto (ítems 1-5), cinco indirectamente relacionadas con ella (ítems 6-10) y cinco de tipo personal, aunque en el ám-

bito del tema de la foto (ítems 11-15).

El test 2 era relativamente similar al 1, si bien en lugar de una foto fija se entregaba al alumno una tira cómica que reproducía una historia, sobre la que se centraba la entrevista.

4. Resultados

Una vez finalizada la transcripción de las entrevistas, se procedió a hacer el recuento por alumno del total de palabras emitidas en inglés en cada entrevista. Los datos resultantes fueron incorporados al programa estadístico SPSS, determinando para cada alumno el tipo de tarea realizada y el centro educativo de procedencia. Una vez dispusimos de todos los datos, se hizo

el análisis de comparación de medias para grupos independientes.

A continuación presentamos las cantidades relativas al número de palabras por alumno en cada una de las tres tareas estudiadas. Los datos correspondientes a la media, la desviación típica y las puntuaciones mínimas y máximas quedan recogidos en el cuadro 2.

Las medias de las producciones orales nos muestran que la entrevista oral a partir de una imagen obtuvo unos resultados muy inferiores a los obtenidos en los otros dos tipos de entrevista. Para averiguar las posibles diferencias significativas entre las medias alcanzadas por estos tres grupos de alumnos, y considerando que estamos trabajando con variables cuantitativas de intervalo o escala, optamos por la prueba estadística que se ajusta a estas limitaciones, el ANOVA de un factor (Peñate *et al.*, 2003: 39), teniendo en cuenta que cuando el ANOVA halla una diferencia significativa entre las medias de varios grupos quiere decir que hay diferencias entre al menos dos de las medias, pero no se indica entre qué medias hay diferencias. Para ello es necesario realizar un análisis posterior (Camacho, 2000: 172). No obstante, el paquete estadístico SPSS nos da la opción de solventar este

problema a través del procedimiento estadístico de contraste de medias dos a dos de Tukey, teniendo además la opción de hacer una prueba de homogeneidad de la varianza (prueba de Levene).

Estos procedimientos estadísticos nos facilitaron los siguientes datos. En este caso, la prueba de Levene (0,481) tiene una probabilidad (0,619) superior a 0,05, lo que viene a indicarnos que no hay diferencias significativas entre las varianzas, es decir, que son similares los tres grupos. Una vez obtenida esta información comprobamos que el test de ANOVA nos indica que hay diferencias significativas entre al menos dos de los tres grupos. Es ahora el procedimiento estadístico de Tukey el que nos señala (cuadro 3, en la página siguiente) que las diferencias entre la tarea de entrevista individual con foto y los otros dos tipos de tareas (entrevista individual con cómic y entrevista en parejas) son significativas, con unos índices de probabilidad altísimos en ambos casos (0,000). Sin embargo, no ocurre lo mismo al comparar la producción oral del alumnado en la tarea individual con el cómic y la tarea en parejas, donde no se aprecia una diferencia significativa (0,992).

Cuadro 2. Media, desviación típica y puntuaciones mínimas y máximas de cada grupo.

	N	Media	Desviación típica	Mínimo	Máximo
Foto - individual	84	165,48	99,113	16	605
Comic - individual	70	226,81	92,133	80	644
Diálogo en parejas	90	224,94	95,894	70	480
Total	244	205,01	99,790	16	644

Cuadro 3. Test de Tukey para la comparación de la media de la producción oral entre los tres grupos de alumnos

Comparaciones múltiples: HSD de Tukey para el total de palabras						
(I) Tarea	(J) Tarea	Diferencia de medias (I-J)	Error típico	Significativa	Intervalo de confianza al 95%	
					Límite inferior	Límite superior
Foto - individual	Comic - individual	-61,338*	15,531	,000	-97,96	-24,71
	Diálogo en parejas	-59,468*	14,559	,000	-93,80	-25,13
Comic - individual	Foto - individual	61,338*	15,531	,000	24,71	97,96
	Diálogo en parejas	1,870	15,293	,992	-34,20	37,94
Diálogo en parejas	Foto - individual	59,468*	14,559	,000	25,13	93,80
	Comic - individual	-1,870	15,293	,992	-37,94	34,20

*. La diferencia de medias es significativa al nivel 0.05.

Conclusiones

A modo de conclusión podemos empezar afirmando que el uso de un tipo u otro de tarea sí tiene incidencia en la cantidad de producción oral del alumnado que hemos estudiado.

Teniendo en cuenta que estamos evaluando un nivel A2, donde aún se presta mayor atención a la cantidad que a la calidad, los datos que nos aporta nuestro primer análisis, sobre el número total de palabras emitidas, deben ser tenidos en cuenta. Es justamente la tarea usada con mayor frecuencia (comentarios sobre una foto) la que

ha dado unas cantidades estadísticamente inferiores a las otras dos tareas. Es decir, el alumnado tuvo mayor dificultad a la hora de hablar sobre lo que podía percibir en la foto. Tal vez nos deberíamos preguntar si dicha dificultad realmente radica en el nivel de la lengua inglesa o en la propia incapacidad del alumnado a la hora de imaginar lo que dicha foto podía representar o qué sucesos podían estar relacionados con la imagen que se les presentó. Y todo ello a pesar de que la foto presentaba una situación (botellón) que le es familiar al alumnado de bachillerato.

En resumen, hemos podido constatar que la tarea de narración a partir del cómic y la entrevista en parejas son las que facilitan una mayor producción oral. Debemos apuntar también que ambas tareas en realidad son muestras de dos de las nue-

vas destrezas redefinidas por el Marco Común Europeo de Referencia para las Lenguas: la expresión y la interacción oral. Destrezas que deben ser implementadas en el proceso de enseñanza-aprendizaje y, por supuesto, en la evaluación.

Referencias bibliográficas

- ARNAIZ, P.; PEÑATE, M. (2004). «El papel de la producción oral (output) en el proceso de aprendizaje de una lengua extranjera (LE)». *Porta Linguarum*, núm. 1, pp. 37-59.
- CAMACHO, J. (2000). *Estadística con SPSS con versión 9 para Windows*. Madrid: RAMA.
- DAVIS, L. (2009). «The influence of interlocutor proficiency in a paired oral assessment». *Language Testing*, vol. 26(3), pp. 367-396.
- FILLMORE, C. (1979). «On fluency». En: FILLMORE, C. (ed.). *Individual differences in language ability and language behaviour* (pp. 85-101). Nueva York: Academic Press.
- NORTON, J. (2005). «The paired format in the Cambridge Speaking Test». *ELT Journal*, vol. 59(4), pp. 287-297.
- PEÑATE, M.; CASTELLANO, A.; BAZO, P. (2003). «El análisis estadístico con programas informáticos». *Lenguaje y Textos*, núm. 21, pp. 37-50.
- ROSS, S.; BERWICK, R. (1992). «The discourse of accommodation in oral proficiency interviews». *Studies in Second Language Acquisition*, vol. 14(2), pp. 159-176.
- SAVILLE, N.; HARGREAVES, P. (1999). «Assessing speaking in the revised FCE». *ELT Journal*, vol. 53(1), pp. 42-51.
- SKEHAN, P.; FOSTER, P. (1999). «The influence of task structure and processing conditions on narrative retellings». *Language Learning*, vol. 49(1), pp. 93-120.
- SWAIN, M. (1985). «Communicative competence: some roles of comprehensible input and comprehensible output in its development». En: GASS, S.; MADDEN, C.G. (eds.). *Input and Second Language Acquisition* (pp. 235-256). Rowley, MA: Newbury House.
- SWAIN, M.; LAPKIN, S. (1995). «Problems in output and the cognitive processes they generate: a step towards second language learning». *Applied Linguistics*, vol. 16(3), pp. 371-391.
- (1998). «Interaction and second language learning: two adolescent French immersion students working together». *The Modern Language Journal*, vol. 82(3), pp. 320-337.
- TAYLOR, L. (2001). «The paired speaking test format: recent Studies». *University of Cambridge ESOL Examinations Research Notes*, núm. 6, pp. 15-17.
- WAGNER-GOUGH, K.; HATCH, E. (1975). «The importance of input in second language acquisition studies». *Language Learning*, vol. 25(2), pp. 297-308.
- YOUNG, R.; MILANOVIC, M. (1992). «Discourse variation in oral proficiency interviews». *Studies in Second Language Acquisition*, vol. 14(4), pp. 403-424.