

ORTOGRAFÍA Y VARIEDAD LÉXICA EN LA PRODUCCIÓN ESCRITA DE LOS APRENDICES ESPAÑOLES DE INGLÉS

ANTONIO R. ROLDÁN TAPIA
Universidad de Córdoba

RESUMEN. *Recientemente parece haberse despertado nuevamente un interés por el producto escrito de los aprendices de inglés. En este trabajo se pretende contribuir al estudio del mismo, a través del análisis de tres elementos del componente léxico de una composición: uno de ellos, ya empleado, la corrección ortográfica, los otros dos, más novedosos, la variedad léxica de nombres y verbos. Los sujetos participantes son estudiantes de COU, de quienes se van a analizar las composiciones hechas del tipo selectividad. Se pretende describir si estos factores léxicos sirven para diferenciar una buena composición escrita, la influencia de los años de instrucción formal en la producción escrita y las diferencias existentes entre aprendices con distinto nivel. Se presenta un análisis casi-experimental, descriptivo cuantitativo, y de correlaciones, previo a las conclusiones.*

PALABRAS CLAVE. *Ortografía, variedad léxica, escritura, aprendiz de L2.*

ABSTRACT. *A renovated interest on the written product of the learners of English has recently come on stage. This paper tries to provide some information by means of the analysis of three lexical components of a written text: one of them, already used, orthography, the other two, new ones, noun and verb variation. Subjects are college-bound secondary school students, whose written compositions are going to be the focus of the analysis. The purpose is threefold: to describe whether these lexical elements characterize a good written composition, to observe the influence of formal instruction on the written output and to distinguish differences in the learners' written performance. Research design is quasi-experimental, with a quantitative descriptive and correlational analysis previous to conclusions.*

KEYWORDS. *Orthography, lexical variation, writing, L2 learner.*

1. INTRODUCCIÓN

Tras unos años en los que la didáctica de la lengua y la investigación sobre el aprendizaje de la misma ha parecido estar orientada mucho más al proceso que al pro-

ducto, por ejemplo en el énfasis puesto en la metodología basada en proyectos (Hutchinson 1991; Ribé & Vidal 1993; Roldán Tapia 1997); ahora parece producirse un movimiento pendular nuevamente hacia el producto de la escritura en L2: por ejemplo, un interés por la ortografía (Madrid y Velásquez 2000), por el efecto del factor edad en el proceso de instrucción (Doiz y Lasagabaster 2001; Torres, Celaya y Pérez-Vidal 2001) por los resultados de pruebas estándar (Sanz Sainz 1999; Chaudron, Martín Ruiz y Whittaker 2001), por las situaciones de bilingüismo entre los aprendices escolares (Celaya 2001; Torres, Celaya y Pérez-Vidal 2001), por la actividad cognitiva que se producen en el proceso de escritura (Manchón, Roca y Murphy 2000), por los mecanismos para medir la habilidad de escritura (Polio 1997; Wolfe-Quintero, Inagaki & Kim 1998) o por revisar el estado de la cuestión en algunos contextos (Reichelt 1999; Tarnopolsky 2000).

En esta situación, este trabajo pretende explorar cuáles son algunos de los elementos que sirven para definir lo que percibimos como una buena composición, a través del análisis del léxico empleado por los aprendices, con la utilización de unas medidas de tipo cuantitativo sobre unas producciones escritas generadas en un contexto de instrucción escolar. Es un intento de cuantificar de forma objetiva lo que habitualmente percibimos como una buena o mala composición escrita en una evaluación de conjunto. Para este propósito se van a analizar algunas de las variables de riqueza léxica que describe la bibliografía especializada.

Las variables de *riqueza léxica* (*lexical richness*) que se pueden medir son las que siguen: Read (2000: 200-201), en su revisión de las investigaciones realizadas, señala la variación léxica, la sofisticación léxica, la densidad léxica y el número de errores. A estas variables, Linnarud (1986: 42-48) añade el número de palabras por oración, el número de palabras por texto y la individualidad léxica. Laufer & Nation (1995) también toman en consideración el perfil de frecuencia léxica (*lexical frequency profile*).

En este trabajo son tres las variables que vamos a medir: la corrección ortográfica, que Linnarud (1986: 50-51) incluye como un tipo concreto de error y la variación léxica en cuanto al uso de nombres y verbos. Si bien no tenemos ninguna referencia anterior a este tipo de medida, -sí la hay para la densidad léxica y para la variedad léxica-, podemos pensar que, como elementos nucleares que son de los sintagmas nominal y verbal, nos podrían dar una indicación más precisa de la utilización del léxico hecha por los aprendices que la que podrían, quizás, ofrecernos adjetivos y adverbios.

En nuestro caso, computamos como errores en la corrección ortográfica palabras inexistentes en inglés (*estudiant**), errores en la base de la palabra (*diferent**) o en los elementos morfológicos (*boxs**). Otras palabras, correctas en cuanto a ortografía, pero utilizadas con otro significado diferente, también se han computado, por ejemplo *bank* en lugar de *bench*.

En la bibliografía especializada (Linnarud 1986; McCarthy 1990; Tribble 1996; Read 2000) se define el concepto de *densidad léxica* (*lexical density*) como la proporción entre el número total de palabras de un texto y el número de *palabras léxicas*

(*lexical words*) existentes en el mismo. La *variedad léxica* (*lexical variation*) es la proporción existente entre el número de palabras totales utilizadas en un texto y el número de palabras diferentes empleadas en el mismo (*type-token ratio*).

En este trabajo se va a utilizar un concepto mixto, a caballo entre ambos, pues vamos a medir la variedad léxica de únicamente dos tipos de palabras por separado: nombres y verbos. Cada nombre o cada verbo se contabilizará una vez independientemente del número de ocasiones en que sea empleado.

Varios son los trabajos que tienen alguna información que aportar a esta investigación. Así, Madrid y Muros (1985) presentan un estudio con alumnos de la antigua E.G.B., correspondiente al actual 6º de Primaria y 1º y 2º de E.S.O., y realizado por medio de dictados preparados previamente. Los alumnos de 6º de E.G.B. hicieron dictados que oscilaban entre 40-60 palabras, con una media de 9 errores cometidos por dictado. Los alumnos de 7º, en dictados de 80 a 90 palabras, cometían una media de 8 errores por dictado. Los de 8º, con textos de hasta 100 palabras, llegaban a la misma media de errores. Transformado en porcentajes esto indica que los alumnos de 6º producían entre un 77.5 y un 85% de palabras correctas, los de 7º entre un 90 y un 91%, mientras que los de 8º llegaban al 92%. En el mismo trabajo se estudió la incidencia que tenía la realización de un dictado previamente preparado frente a uno completamente nuevo. La reducción en el porcentaje de errores tras el dictado preparado está entre un 52% de los alumnos de 7º, y un 42% de los de 8º. Los estudiantes de Magisterio se sometieron al mismo tipo de prueba, observándose una reducción de errores del 49%.

Madrid y Velásquez (2000) realizan un estudio de la ortografía en la producción escrita de alumnos del 2º ciclo de Primaria, en la que los sujetos participantes son dos grupos: uno de 32 alumnos de 3º curso y otro de 33 alumnos de 4º curso, de un centro de procedencia urbana. En el apartado de conclusiones, a esa edad y tras aproximadamente 120 horas de instrucción, el alumnado ha fijado su ortografía en un 55% y un 62%, respectivamente, en las pruebas de reconocimiento visual, disminuyendo ese porcentaje al 32% y 34% en las pruebas de percepción auditiva.

Doiz y Lasagabaster (2001) al comparar el factor edad a igual cantidad de tiempo de instrucción utilizan tres grandes bloques de medidas –fluidez, complejidad y precisión- con una serie de variables incluidas, que resultan de interés en esta revisión. Los sujetos de su investigación son 30 alumnos de 2º de ESO y 30 alumnos de 1º de Bachillerato, pertenecientes al mismo centro escolar. Aunque algunas de las variables necesitan cierta aclaración, se observa que los alumnos de Bachillerato, de mayor edad, utilizan mayor número de verbos no-finitos, mayor número de verbos léxicos y mayor número de auxiliares. En cuanto al porcentaje de errores ortográficos, los alumnos de 2º de ESO, tienen una media del 6,29%, que es igual a un 93.71% de corrección; mientras que los de Bachillerato obtienen un 2,70%, que supone un 97.3% de corrección.

El estudio de Chaudron, Martín y Whittaker (2001) sobre las composiciones tipo selectividad demuestra cómo ciertos tratamientos específicos en la metodología de la

composición mejoran en términos generales la fluidez, complejidad y corrección de las producciones escritas de los alumnos. Éstas varían entre 80.6 palabras de media como mínimo y 154.8 como máximo. En el mismo estudio se recoge el dato de que el 60% de los estudiantes examinados en la Universidad Carlos III no superan el 50% de la calificación posible de la composición.

El trabajo de Linnarud (1986) es interesante porque se trata de estudiantes de unas características muy parecidas a los sujetos de esta investigación, compartiendo con ellos las siguientes: (i) son estudiantes de inglés como lengua extranjera, (ii) tienen una media de edad de 17 años, (iii) son estudiantes de enseñanza secundaria post-obligatoria y de la opción de Humanidades, (iv) han estudiado hasta el momento de la investigación aproximadamente unas 640 horas de inglés en el sistema educativo, (v) tienen que hacer un examen normalizado, al igual que los estudiantes españoles hacen Selectividad, y (vi) son alumnos de procedencia urbana. Aunque la metodología y los objetivos de su investigación difieren –compara 42 textos en inglés de estudiantes suecos con 21 textos de estudiantes nativos, con la participación de tres grupos de evaluadores externos: profesores suecos de inglés, profesores nativos universitarios y hablantes nativos de inglés-, una parte importante de los factores lingüísticos (Linnarud 1986: 42-48) que analiza en las composiciones son semejantes a los de este estudio. En el análisis del número de errores ortográficos por composición, solamente 6 de las 42 composiciones no tienen ningún error de este tipo, situándose la media en 1.8 y la desviación estándar en 1.46. Los resultados de la variación léxica demuestran que los sujetos aprendices del estudio varían mucho menos que los hablantes nativos del mismo estudio, con una media del 63.6 entre los primeros y del 70.8 entre los segundos. Los resultados de la densidad léxica son semejantes a los de la variedad léxica, con la diferencia de que las medias para cada grupo no son tan distantes una de la otra: 42.4 para los sujetos aprendices y 44.0 para los nativos. De la correlación de los datos de los sujetos aprendices se desprende que hay más errores y repetición léxica cuanto más larga es la composición y, además, se observa que cuanto mayor es el número de errores, mayor es la repetición de las mismas palabras en la composición. De la comparación entre los datos obtenidos de las composiciones y las evaluaciones hechas por los tres grupos de evaluadores, se concluye que (a) hay una diferencia estadísticamente significativa entre el número de palabras producidas por los sujetos aprendices y por los nativos; (b) que también existe correlación entre el número de palabras y la evaluación dada a la composición, si bien no existe relación directa entre una composición extensa y una buena puntuación, sí se produce entre una composición breve y una baja puntuación; (c) hay una enorme diferencia en el número de errores producidos entre aprendices y nativos, si bien los pocos errores cometidos por los nativos siempre tienen que ver con la ortografía; (d) se produce también una correlación positiva entre la variación léxica y la puntuación obtenida por cada composición, tanto entre las peores como entre las mejores; y (e) no se observa correlación significativa entre la densidad léxica y las puntuaciones dadas a cada composición.

2. DISEÑO DE LA INVESTIGACIÓN

2.1. *Objetivos*

Intentar explorar qué factores de tipo léxico pueden contribuir a la producción de una buena composición escrita realizada por un estudiante de enseñanza secundaria, preuniversitaria. Para ello, se seleccionan ciertos elementos como la ortografía y la variedad léxica de nombres y verbos del texto.

Sondear cuál es la influencia del número de horas de instrucción formal en inglés en la producción léxica de estos mismos estudiantes de secundaria.

2.2. *Sujetos*

Los sujetos participantes son 6 estudiantes de COU, de un centro público, de extracción urbana, procedentes de una zona de la ciudad cuya situación socioeconómica podría definirse como favorable y en crecimiento. La media de edad de estos sujetos se haya entre los 17 y 19 años, y cuentan con un mínimo de 640 horas de instrucción en el sistema escolar. La opción de COU que todos hacen es la de Humanidades.

Para el propósito de la investigación los seis estudiantes se dividen en dos grupos claramente definidos. El grupo 1 lo forman estudiantes que en las últimas 5 evaluaciones han obtenido sobresaliente o notable como calificación (es decir, entre un 100-75% de la puntuación posible), mientras que el grupo 2 lo forman estudiantes que han obtenido una calificación de insuficiente o suficiente (entre un 35-55% de su posible puntuación total), en el mismo número que el grupo 1, es decir 3.

Los tres estudiantes del grupo 1 serán llamados S1, S2 y S3, mientras que los del grupo 2, serán, en todo momento S4, S5 y S6. El grupo 1 lo forman tres sujetos de sexo femenino, mientras que en el grupo 2, S4 y S6 son de sexo masculino.

2.3. *Metodología*

En primer lugar, como cualquier otra investigación generada a partir de situaciones existentes en el mundo real (Hatch & Farhady 1982: 22-23; Seliger & Shohamy 1989: 148), debemos señalar que se trata de un diseño de investigación *casi-experimental* (*quasi-experimental*), representativo de las condiciones naturales que se dan en el contexto educativo.

Para todo el estudio se han tomado como instrumentos de la investigación las composiciones tipo-selectividad producidas por los alumnos. Estos textos, cuya extensión está perfectamente regulada por una normativa entre 80 y 100 palabras, corresponden a la última pregunta de la prueba, tras haber realizado previamente las de comprensión, gramática y vocabulario. Se debe aclarar que en un buen número de casos, las composiciones sobrepasan el límite de 100 palabras; a pesar de ello, se aceptan para la investigación.

Como paso previo, se han valorado tres series de composiciones realizadas por los alumnos, como prueba piloto (Hatch & Farhady 1982: 36; Seliger & Shohamy 1989: 195), para comprobar las variables que posteriormente vamos a medir. Además, en este

mismo momento de la investigación, se ha revisado el concepto de palabra, en especial en aquellos casos en los que no hay una correspondencia entre un significado y una palabra aislada, produciéndose *mutliword units*, que Schmitt (2000: 2) define como “*an item that functions as a single meaning unit, regardless of the number of words it contains*”.

Una vez realizado el paso anterior, se han tomado tres composiciones de cada alumno, entre las que éstos han hecho en casa como tarea normal y propia del curso. A cada composición se le ha contabilizado el número total de palabras producidas, el porcentaje de palabras cuya ortografía es correcta, el número de nombres diferentes empleados, así como los verbos también diferentes utilizados. Cada una de estas tres composiciones se denominará como T1, T2 y T3 a lo largo de toda la investigación. Los temas para estas tres composiciones son:

T1: *Art and new technologies*

T2: *Hollywood and movie industry*

T3: *Animal lovers*

Se han tomado dos composiciones más de cada alumno, como control, tomadas de los exámenes propios de evaluación, para comprobar si la calidad de la producción escrita de cada alumno se repite en todas las circunstancias. A cada una de estas composiciones de control, se las denominará como C1 y C2 de aquí en adelante. No existe un límite de tiempo para su realización, al igual que para las composiciones T. Los temas propuestos para estas composiciones son:

C1: *Shopping at Christmas*

C2: *School violence*

Para T1, T2 y T3 se medirán la media de cada alumno y la media del grupo, al igual que para C1 y C2. Así mismo, se medirá la diferencia entre las medias de los textos T1, T2 y T3 y la de los C1 y C2.

Como último paso, se correlacionarán las medias de todas las variables con el objeto de buscar ciertos patrones significativos en los valores de la producción escrita de los aprendices.

Se ha intentado controlar en todo momento un buen número de variables que pudiesen afectar los resultados: así, (a) todos los aprendices han recibido el mismo tipo de instrucción, que se resume en unas 640 horas de instrucción en clase, con unos programas marcadamente comunicativos en los primeros cursos y un énfasis en las destrezas de lectura y escritura y gramática en el último año; (b) la longitud de los textos está claramente delimitada por la normativa que regula la prueba de Selectividad, aunque puede haber alguna variación al alza; (c) el tipo de tarea escrita es la misma; (d) los temas de las composiciones, indicados anteriormente, son además iguales para todos ellos, y (e) el factor tiempo no condiciona las composiciones C, pues a pesar de estar hechas en situación de examen, la duración máxima de la prueba con la que han contado los estudiantes para realizar las C1 y C2 es tal que duplica el tiempo regulado para la Selectividad, permitiéndoles la composición a un ritmo adecuado a sus capacidades. De

hecho, todos los sujetos concluyeron sus composiciones bastante antes del límite máximo permitido.

3. RESULTADOS Y DISCUSIÓN

Se presentan a continuación dos tipos de datos y análisis: el primero de ellos, de datos cuantificados, con medias y porcentajes, así como las diferencias entre las composiciones T y C; el segundo, un análisis estadístico de correlaciones de los datos de todas las variables empleadas.

3.1. *Análisis descriptivo*

NÚMERO DE PALABRAS DE COMPOSICIONES "T"				
	T1	T2	T3	M
S1	95	105	88	96
S2	91	101	91	94.33
S3	88	97	82	89
S4	122	126	82	110
S5	92	102	89	94.33
S6	105	84	76	88.33

Tabla 1. *Número de palabras en composiciones "T"*.

Se observa que todos los sujetos producen composiciones con un número de palabras acorde con la norma sobre textos de Selectividad, que regula el número entre 80-100, a excepción, claramente del sujeto S4. En su conjunto los estudiantes del Grupo 2 (S4-S5-S6) producen textos más extensos que los tres estudiantes del Grupo 1. Este dato se asemeja a una de las conclusiones aportadas por Linnarud (1986) en cuanto a que no hay una correlación entre el número de palabras de una composición y la evaluación dada a la misma. En nuestro estudio, precisamente, las composiciones más extensas son producidas por estudiantes con más baja calificación. Por último, cabe señalar que todos los estudiantes, excepto el S3, superan, por defecto o por exceso, al menos en una ocasión, el número de 80-100 palabras.

NÚMERO DE PALABRAS COMPOSICIONES "C"			
	C1	C2	M
S1	115	109	112
S2	162	121	141

NÚMERO DE PALABRAS COMPOSICIONES "C"			
S3	93	82	87.5
S4	105	116	110.5
S5	104	100	102
S6	87	82	84.5

Tabla 2. *Número de palabras en composiciones "C"*.

La media (M) de palabras producidas por cada estudiante en las composiciones C supera a las producidas en las composiciones T, a excepción de los estudiantes S3 y S6. Del mismo modo que antes, todos los sujetos superan la media de 80-100 palabras, todos por exceso, menos los estudiantes, nuevamente, S3 y S6. Destaca sobremanera la C1 del estudiante S2 con 162 palabras. Por tanto, se puede pensar que en condición de examen, la ansiedad que esta situación genera, hace aumentar el número de palabras de la composición. Este margen entre 162 palabras máximo y 82 mínimo coincide con los datos presentados por Chaudron, Martín y Whittaker (2001), que oscilan entre 80.6 y 154.8 de medias mínima y máxima respectivamente.

PORCENTAJE DE CORRECCIÓN ORTOGRÁFICA COMPOSICIONES "T"				
	T1	T2	T3	M
S1	98.94	99.04	100	99.32
S2	97.80	97.02	97.80	97.54
S3	97.72	98.96	98.78	98.48
S4	97.54	95.23	97.56	96.77
S5	97.82	95.09	96.62	96.51
S6	96.19	97.61	97.36	97.05

Tabla 3. *Porcentaje de corrección ortográfica en composiciones "T"*.

Se observa un alto porcentaje de corrección ortográfica en ambos grupos, G1 y G2, con un 100% en la T3 de S1 como porcentaje más alto y un 95.23% en la T2 de S4 como más bajo. Se observa, en cualquier caso, mayor corrección en el G1, con aproximadamente casi 2 puntos de diferencia, con respecto a los sujetos de G2. Comparativamente, se aprecia el incremento en la corrección con respecto al 62% de máxima en los alumnos de 2º de Primaria (Madrid y Velásquez 2000) y un 92% de los alumnos de 8º de EGB (Madrid y Muros 1985). Lo mismo ocurre con el 93.71% de los alumnos de 2º ESO y el 97.3% de los de Bachillerato (Doiz y Lasagabaster 2001). El 1.8 de errores de las com-

posiciones de los estudiantes suecos analizadas por Linnarud (1986) supone un 98.2% de corrección, equivalente a las medias obtenidas por los sujetos de G1.

PORCENTAJE DE CORRECCIÓN ORTOGRÁFICA COMPOSICIONES "C"			
	C1	C2	M
S1	99.13	98.16	98.64
S2	98.76	97.52	98.14
S3	98.92	96.34	97.63
S4	90.47	96.55	93.51
S5	94.23	97.00	95.61
S6	94.25	92.68	93.46

Tabla 4. *Porcentaje de corrección ortográfica en composiciones "C"*.

Se repite, igualmente, el alto porcentaje de corrección en ambos grupos, con mayor índice de corrección en G1. En términos generales, la media es menor en prácticamente todos los sujetos, con una tendencia a la baja en los sujetos de G2. El efecto de la ansiedad de la situación de examen puede explicar esta diferencia de medias. Las medias de control del G1 siguen estando en torno al 98.2% de los estudiantes suecos, mientras que las medias de control del G2 bajan, en especial S4 y S6, acercándose al 92% de los alumnos de 8º de EGB (Madrid y Muros 1985).

VARIEDAD LÉXICA DE NOMBRES COMPOSICIONES "T"				
	T1	T2	T3	M
S1	15.78	18.09	13.63	15.83
S2	19.78	18.81	16.48	18.35
S3	23.86	13.40	15.85	17.70
S4	14.75	11.11	14.63	13.49
S5	11.95	13.72	12.35	12.67
S6	14.28	13.09	13.15	13.50

Tabla 5. *Variedad léxica de nombres en composiciones "T"*.

Las medias de G1 y G2 señalan importantes diferencias, pues entre la media más baja de G1, 15.83% de S1, y la más alta de G2, 13.50% de S6, hay más de 2 puntos de diferencia. Esta diferencia se incrementa si tomamos la media más alta, en T1 de S3, con

un 23.86%, y la más baja, en T2 de S4, con un 11.11%, con una diferencia de más de 12 puntos entre una y otra.

VARIEDAD LÉXICA DE NOMBRES COMPOSICIONES "C"			
	C1	C2	M
S1	13.91	16.51	15.21
S2	13.58	12.39	12.98
S3	12.90	15.85	14.37
S4	12.38	13.79	13.08
S5	13.46	14.00	13.73
S6	13.79	10.97	12.38

Tabla 6. *Variedad léxica de nombres en composiciones "C"*.

Estas composiciones de control vuelven a repetir el mismo patrón de funcionamiento: mejores medias en sujetos de G1 que en sujetos de G2 y tendencia a la baja en prácticamente todas las medias. Se puede volver a apuntar el efecto de ansiedad en situación de examen. La diferencia entre la media más alta, con un 16.51% de S1 en C2 y la media más baja, con un 10.97% de S6 en C2, está en 5.54 puntos. Si se contrasta con las medias de las composiciones T este margen de diferencia es mucho más reducido: de 12.75 puntos en aquéllas y tan sólo 5.54 en éstas.

VARIEDAD LÉXICA DE VERBOS COMPOSICIONES "T"				
	T1	T2	T3	M
S1	9.47	9.52	17.04	12.01
S2	14.28	8.91	12.08	11.75
S3	9.09	11.34	9.87	10.10
S4	8.19	7.93	7.31	7.81
S5	9.78	10.78	7.86	9.47
S6	5.71	8.33	11.84	8.62

Tabla 7. *Variedad léxica de verbos en composiciones "T"*.

Se observa una repetición en unas medias mejores en el caso de los sujetos S1, S2 y S3 frente a las medias de S4, S5 y S6. En términos generales las medias son todas más bajas comparadas con las de los nombres. Hay diferencias muy destacables entre la

media más alta, 17.04% de S1 en T3, y la más baja, 5.71 de S6 en T1, con un margen de diferencia entre ambas de 11.33 puntos.

VARIEDAD LÉXICA DE VERBOS COMPOSICIONES "C"			
	C1	C2	M
S1	12.06	11.01	11.53
S2	7.40	4.95	6.17
S3	9.67	9.75	9.71
S4	5.71	5.17	5.44
S5	5.76	13.00	9.38
S6	10.34	6.09	8.21

Tabla 8. *Variedad léxica de verbos en composiciones "C"*.

Estas composiciones C vuelven a repetir valores semejantes a los que ofrecían para los nombres: (a) medias más altas en sujetos del G1 que del G2; (b) tendencia a la baja en todos los sujetos; y (c) reducción de la distancia entre la media más alta y más baja, 12.06% de S1 en C1 y 5.17 de S4 en C2: 5.89 puntos de diferencia, casi la mitad que en las composiciones T.

SUJETOS	N. PALABRAS			ORTOGRAFÍA			VAR. NOMBRES			VAR. VERBOS		
	M-T	M-C	DIF.	M-T	M-C	DIF.	M-T	M-C	DIF.	M-C	M-T	DIF.
S1	96	112	16	99.32	98.64	-0.68	15.83	15.21	-0.62	12.01	11.53	-0.48
S2	94.33	141.54	7.17	97.54	98.14	0.60	18.35	12.98	-5.37	11.75	6.17	-5.58
S3	89	87.5	-1.5	98.48	97.63	-0.85	17.70	14.37	-3.33	10.10	9.71	-0.39
MG1	93.11	113.66	20.55	98.44	98.13	-0.31	17.29	14.18	-3.11	11.28	9.13	-2.15
S4	110	110.50	0.50	96.77	93.51	-3.26	13.49	13.08	-0.41	7.81	5.44	-2.37
S5	94.33	102	7.67	96.51	95.61	-0.90	12.67	13.73	1.06	9.47	9.38	-0.09
S6	88.33	84.5	-3.83	97.05	93.46	-3.59	13.50	12.38	-1.12	8.62	8.21	-0.41
MG2	97.55	99	1.45	96.77	94.19	-2.58	13.22	13.06	-0.16	8.63	7.67	-0.96

Tabla 9. *Medias de sujetos y grupos.*

De la comparación de todos los datos, se observa que mientras las medias de ambos grupos suben en la producción del número de palabras en las composiciones C, en el resto de valores –ortografía y variedad léxica- siempre tienden a la baja en C.

Los sujetos del G1 presentan poca diferencia entre las medias de sus composiciones T y C en lo que respecta a la ortografía, mientras que los del G2 ven disminuida su media en más de 2 puntos. Esta diferencia se invierte en la variedad léxica de nombres

y verbos, en la que los sujetos de G1 ven disminuir las medias de sus composiciones C en mucha mayor medida que los sujetos del G2. Comparativamente, la variedad léxica de nombres supera a la de verbos en las medias obtenidas por todos los sujetos, tanto en las composiciones T como en C.

Prestando atención a las diferencias entre las medias de T y C en ortografía y variedad de nombres y verbos, en el G1 el sujeto más regular es S1, seguido de S3 y S2, con diferencias importantes entre ambos y el primero. En el G2, el estudiante más regular es S5, más incluso que S1, seguido de S6 y S4, con pequeña diferencia entre ambos, aunque considerable con respecto al primero.

3.2. Análisis de correlaciones

		VAR01	VAR02	VAR03	VAR04	VAR05	VAR06	VAR07	VAR08	VAR09
VAR01	Correlación Pearson Sig. (bilateral)	1,000	.311 .549	-.387 .448	-.841 .036	-.932 .007	-.930 .007	-.595 .213	-.866 .026	-.349 .498
VAR02	Correlación Pearson Sig. (bilateral)	.311 .549	1,000	.382 .455	-.272 .602	-.320 .536	-.311 .549	-.059 .911	-.350 .497	-.524 .286
VAR03	Correlación Pearson Sig. (bilateral)	-.387 .448	.382 .455	1,000	.039 .941	.443 .379	.434 .390	-.028 .958	.543 .265	-.412 .417
VAR04	Correlación Pearson Sig. (bilateral)	-.841 .036	-.272 .602	.039 .941	1,000	.772 .072	.629 .181	.790 .061	.723 .105	.655 .158
VAR05	Correlación Pearson Sig. (bilateral)	-.932 .007	-.320 .536	.443 .379	.772 .072	1,000	.791 .061	.710 .114	.947 .004	.493 .320
VAR06	Correlación Pearson Sig. (bilateral)	-.930 .007	-.311 .549	.434 .390	.629 .181	.791 .061	1,000	.291 .575	.719 .107	.041 .939
VAR07	Correlación Pearson Sig. (bilateral)	-.595 .213	-.059 .911	-.028 .958	.790 .061	.710 .114	.291 .575	1,000	.575 .233	.788 .063
VAR08	Correlación Pearson Sig. (bilateral)	-.866 .026	-.350 .497	.543 .265	.723 .105	.947 .004	.719 .107	.575 .233	1,000	.456 .365
VAR09	Correlación Pearson Sig. (bilateral)	-.349 .498	-.524 .286	-.412 .417	.655 .158	.493 .320	.041 .939	.788 .063	.455 .365	1,000

Tabla 10. *Correlaciones entre variables.*

Variables:

VAR01: Grupos G1 y G2.

VAR02: Número de palabras en composiciones T

VAR03: Número de palabras en composiciones C

VAR04: Corrección ortográfica en composiciones T

VAR05: Corrección ortográfica en composiciones C

VAR06: Variedad léxica de nombres en composiciones T

VAR07: Variedad léxica de nombres en composiciones C

VAR08: Variedad léxica de verbos en composiciones T

VAR09: Variedad léxica de verbos en composiciones C

Los valores de la “r” de Pearson nos muestran una serie de correlaciones negativas, de valor cercano a -1, entre los sujetos participantes y las medias de corrección ortográfica en T (-.841), de corrección ortográfica en C (-.932), de variedad léxica de nombres en T (-.930) y de variedad léxica de verbos en T (-.866).

Esta correlación negativa significa que decrecen los valores de corrección ortográfica T y C, y la variedad de nombres y verbos en T, a la vez que decrece la calificación de los sujetos participantes; es decir, los valores de las variables son más bajos para los sujetos S4, S5 y S6 que para los sujetos S1, S2 y S3.

Cabe señalar que únicamente se produce correlación negativa en T y C en la corrección ortográfica, lo que viene a reafirmar que la ortografía es un elemento del léxico muy consolidado en la interlengua de los aprendices en esta fase de aprendizaje.

4. CONCLUSIONES.

De todo lo anterior, y teniendo en cuenta que de forma *quasi-experimental* consideramos que los estudiantes con mejor calificación escolar tienen un mejor nivel de conocimiento de la lengua, se puede concluir lo siguiente:

- (i) no se puede afirmar que exista una relación directa entre el número de palabras de una composición y la riqueza léxica de la misma;
- (ii) en lo que respecta al diseño de la tarea escrita, resulta muy difícil conseguir que el aprendiz de una lengua ajuste el desarrollo de un tema propuesto en un margen, en cuanto a número de palabras, tan acotado y estricto, como son 80-100;
- (iii) la corrección ortográfica es un elemento léxico muy consolidado, con un alto porcentaje de consecución, incluso tras sólo varios años de instrucción escolar, que a este nivel sufre muy pequeña variación independientemente de las condiciones en las que se realiza la composición;
- (iv) los condicionantes externos, -como puede ser una situación de examen-, afectan a algunos elementos del componente léxico de la producción escrita de los aprendices de una lengua;
- (v) los alumnos cuyo nivel de aprendizaje es más elevado presentan mayor corrección ortográfica, y mayor conocimiento del vocabulario, al menos en lo que respecta a nombres y verbos. De lo cual se puede afirmar que estos tres criterios se pueden tomar como medidas de riqueza léxica para distinguir entre aprendices de la lengua;
- (vi) existe una correlación importante entre los tres tipos de variables siempre que no haya un condicionante externo que afecte al proceso de producción escrita.

BIBLIOGRAFÍA.

Celaya, M. L. 2001. “I’ve got 12 years old: L1 in SLA”. *APAC of News* (Bulletin de l’Associació de Professors d’Anglès de Catalunya) 42: 42-48.

- Chaudron, C., Martín Ruíz, A. y R. Whittaker. 2001. "La composición como comunicación: influencia en el desarrollo general del inglés como segunda lengua en un contexto de instrucción". *La Lingüística Aplicada a finales del siglo XX. Ensayos y propuestas*. Tomo 1. Eds. I. de la Cruz, C. Santamaría, C. Tejedor, y C. Valero. Alcalá de Henares: Universidad de Alcalá de Henares y AESLA. 53-62.
- Doiz, A. y D. Lasagabaster. 2001. "El efecto del factor edad en la producción escrita en inglés". *La Lingüística Aplicada a finales del siglo XX. Ensayos y propuestas*. Tomo 1. Eds. I. de la Cruz, C. Santamaría, C. Tejedor y C. Valero. Alcalá de Henares: Universidad de Alcalá de Henares y AESLA. 63-68.
- Hatch, E. & H. Farhady. 1982. *Research Design and Statistics for Applied Linguistics*. Rowley, Mass.: Newbury House.
- Hutchinson, T. 1991. *Introduction to Project Work*. Oxford: Oxford University Press.
- Laufer, B. & P. Nation. 1995. "Vocabulary size and use: lexical richness in L2 written production". *Applied Linguistics* 16/3: 307-322.
- Linnarud, M. 1986. *Lexis in Composition: a performance analysis of Swedish learners' written English*. Malmö: CWK Gleerup.
- Madrid, D. y J. Muros. 1985. "La ortografía del inglés: análisis y tratamiento de las deficiencias ortográficas en alumnos de E.G.B. y Magisterio". *Actas del II Congreso Nacional de Lingüística Aplicada*. Granada: AESLA. 107-120.
- Madrid, D. y E. Velásquez. 2000. "La ortografía del inglés en el 2º Ciclo de Primaria". *GRETA, Revista para Profesores de Inglés* 8/2: 5-14.
- Manchón, R. M., Roca, J. y L. Murphy. 2000. "La influencia de la variable «grado de dominio de la L2» en los procesos de composición en lengua extranjera: hallazgos recientes de la investigación". *Segundas lenguas. Adquisición en el aula*. Ed. C. Muñoz. Barcelona: Ariel. 277-297.
- McCarthy, M. 1990. *Vocabulary*. Oxford: Oxford University Press.
- Polio, C. G. 1997. "Measures of linguistic accuracy in second language writing research". *Language Learning* 47/1: 101-143.
- Read, J. 2000. *Assessing Vocabulary*. Cambridge: Cambridge University Press.
- Reichelt, M. 1999. "Toward more comprehensive view of L2 writing: foreign language writing in the U.S.". *Journal of Second Language Writing* 8/2: 181-204.
- Ribé, R. & N. Vidal. 1993. *Project Work, Step by Step*. Oxford: Heinemann.
- Roldán Tapia, A. R. 1997. *La integración de un "process syllabus" en un curso de Secundaria: input, interacción, destrezas, L1 y L2 en la realización de "project work"*. Granada: Servicio de Publicaciones de la Universidad.
- Sanz Sainz, I. 1999. "El examen de Selectividad a examen". *GRETA, Revista para Profesores de Inglés* 7/2: 16-29.
- Schmitt, N. 2000. *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press.
- Seliger, H. W. & E. Shohamy. 1989. *Second Language Research Methods*. Oxford: Oxford University Press.

- Tarnopolsky, O. 2000. "Writing English as a foreign language: a report from Ukraine". *Journal of Second Language Writing* 9/3: 209-226.
- Torres, R. M., Celaya, M. L. & C. Pérez-Vidal. 2001. "Written performance in EFL by Catalan/Spanish bilinguals". *Trabajos en Lingüística Aplicada*. Coord.. C. Muñoz. Barcelona: AESLA. 223-229.
- Tribble, C. 1996. *Writing*. Oxford: Oxford University Press.
- Wolfe-Quintero, K., Inagaki, S. & H. Kim. 1998. *Second Language Development in Writing: measures of fluency, accuracy and complexity*. Honolulu, SLTCC: University of Hawaii.