

Alicia San-Mateo Valdehíta y
M.ª Antonieta Andión Herrero

DOI:10.4312/vh.27.1.209-229

Universidad Nacional de Educación a Distancia



Eficacia de tres actividades de aprendizaje de vocabulario en español LE/L2: la selección de definiciones, la selección de ejemplos y la escritura de oraciones¹

Palabras clave: español segunda lengua/lengua extranjera, aprendizaje de vocabulario, actividades de aprendizaje, Hipótesis del Nivel de Participación o Implicación, Técnica de Análisis de Características.

1 Introducción

El dominio del vocabulario en una segunda lengua (L2) es una de las competencias básicas del aprendiz, tal y como establece el *Marco común europeo de referencia para las lenguas (MCER)* (Consejo de Europa, 2002); de ahí la relevancia de las investigaciones sobre actividades de aprendizaje del léxico, la eficacia de las mismas y el conocimiento que conllevan.

En este artículo se presentan los resultados globales del proyecto «Grados de eficacia en ejercicios de incorporación de vocablos al lexicon de aprendices de español segunda lengua/lengua extranjera» (FFI2013-44117-P), cuyo objetivo ha sido la aplicación de un riguroso experimento a 1317 estudiantes de español fundamentalmente de nivel B1-B2, jóvenes y con dominio de otra lengua, hombres y mujeres, de 27 lenguas maternas (francés, portugués, italiano, inglés, alemán...), en contexto de lengua extranjera (LE) (Brasil, Estados Unidos de América, Francia e Italia) y de L2 (España).

1 Este estudio ha sido financiado por el proyecto «Grados de eficacia en ejercicios de incorporación de vocablos al lexicon de aprendices de español segunda lengua/lengua extranjera» (FFI2013-44117-P), del Programa Estatal de Investigación Científica y Técnica de Excelencia, del Ministerio de Economía y Competitividad de España.

Tras un breve marco teórico sobre lo más relevante para nuestra investigación, describimos a los participantes, materiales y procedimiento del experimento que persigue determinar la eficacia en el aprendizaje léxico de tres actividades (selección de definiciones, de ejemplos y escritura de oraciones), apoyándonos en las escalas de la *Hipótesis del Nivel de Participación* (Hulstijn y Laufer, 2001) y de la *Técnica de Análisis de Características* (Nation y Webb, 2011), y en el conocimiento, receptivo o productivo, que se adquiere con cada una de ellas.

2 Breve marco teórico

Según Nation (2001: 23-59), conocer una palabra supone tener en cuenta 9 aspectos, agrupados en tres categorías: Forma, Significado y Uso, desde el punto de vista del conocimiento receptivo –o de reconocimiento– y productivo –o de uso–. La Forma incluye (1) la pronunciación, (2) la escritura y (3) las partes de la palabra. El Significado incluye (4) la conexión entre forma y significado, (5) conceptos y referentes, y (6) asociaciones. El Uso incluye (7) funciones gramaticales, (8) colocaciones y (9) restricciones de uso (frecuencia, registro...).

Hulstijn y Laufer (2001) establecen una relación entre la atención que el aprendiz presta a los aspectos que conforman el conocimiento de la unidad léxica y la probabilidad de aprenderla, y establecen una escala, gracias a la *Hipótesis del Nivel de Participación o Implicación (HNP)* (2001: 543-545), que permite clasificar las actividades por su demanda de esfuerzo cognitivo. Emplean tres factores: *Necesidad* (motivación), *Búsqueda* de la palabra (forma) o del significado y *Evaluación* sobre la idoneidad de la palabra en el contexto. Si el factor está presente en la actividad de forma *moderada*, recibe 1 punto y 2 puntos si es mayor (intensidad *fuerte*). Cuanto más alta sea la puntuación, mayor será el nivel de participación e implicación de la tarea y, por lo tanto, habrá más posibilidades de que la palabra se incorpore al léxico mental.

Para buscar un modelo más explicativo, Nation y Webb (2011) proponen la *Técnica de Análisis de Características (TAC)*, basada en 5 factores: *Motivación*, *Percepción o reconocimiento*, *Recuperación*, *Generación* y *Retención*, y en una serie de preguntas para determinar su presencia y grado (2011: 7):

(1) *Motivation:*

1. *Is there a clear vocabulary learning goal?*
2. *Does the activity motivate learning?*
3. *Do the learners select the words?*

(2) *Noticing*

4. *Does the activity focus attention on the target words?*
5. *Does the activity raise awareness of new vocabulary learning?*
6. *Does the activity involve negotiation?*

(3) *Retrieval*

7. *Does the activity involve retrieval of the word?*
8. *Is it productive retrieval?*
9. *Is it recall?*
10. *Are there multiple retrievals of each word?*
11. *Is there spacing between retrievals?*

(4) *Generation*

12. *Does the activity involve generative use?*
13. *Is it productive?*
14. *Is there a marked change that involves the use of other words?*

(5) *Retention*

15. *Does the activity ensure successful linking of form and meaning?*
16. *Does the activity involve instantiation?*
17. *Does the activity involve imaging?*
18. *Does the activity avoid interference?*

Solo cuando la respuesta es afirmativa, la actividad recibe un punto; las de mayor puntuación serán las más eficaces.

Esta escala no ha sido muy aplicada hasta ahora en el campo de la enseñanza del español quizás por la complejidad que supone atender un mayor número de factores. Zou, Wang, Kwan y Xie (2018) evalúan la capacidad de predicción de la escala con tres tipos de actividades (entre ellas, la de escritura de oraciones) y 120 estudiantes de inglés como L2; y confirman la validez del modelo. Por otro lado, en los trabajos de Chaharlang y Farvardin (2018), Gohar, Rahmanián y Soleimani (2018), Hu y Nassaji (2016), y Khoshsima y Eskandari (2017), llevados a cabo con aprendices de inglés como L2, se comparan los dos marcos teóricos y la conclusión en todos ellos es que la TAC es más eficaz

que la *HNP* al predecir los efectos de las diversas actividades de aprendizaje de vocabulario.

En cuanto a qué actividades son más eficaces para aprender vocabulario en una L2, el interés sigue existiendo: “No wonder, then, that both researchers and language teachers have been searching for effective methods to facilitate the acquisition of thousands of words that are necessary to narrow the gap between the vocabulary learners know and the vocabulary they need” (Laufer, 2017: 5), si bien es difícil comparar los resultados de los diferentes estudios porque emplean escalas y sistemas de puntuación diversos, como señala Laufer (2017: 9). De ahí que hasta ahora no haya conclusiones definitivas, sobre todo en lo que al español se refiere, y el debate sobre la efectividad de las actividades siga abierto.

Para dar la mayor validez posible a nuestro estudio, hemos aplicado las dos escalas (*HNP* y *TAC*) a las actividades propuestas para comprobar también su predicción de la eficacia de las actividades de aprendizaje.

3 El presente estudio

Este estudio es una réplica adaptada y enriquecida del que Coomber, Ramstad y Sheets (1986) llevaron a cabo en el ámbito de la lengua materna (inglés) y después Matanzo (1991) y Reyes (1995) (español), y cuya conclusión fue que la actividad más eficaz para aprender nuevas palabras era la escritura. Posteriormente, San-Mateo (2005, 2012) replicó dos veces la investigación en el español como L2 con una muestra de hablantes de inglés. Sus resultados no pueden considerarse definitivos, dado el número reducido de participantes, de ahí el interés de nuestro proyecto (FFI2013-44117-P) en ampliar la investigación con un número elevado de informantes, diversas lenguas maternas y diferentes contextos de aprendizaje².

Nuestro trabajo persigue varios objetivos:

- I. Analizar la eficacia de tres actividades de aprendizaje de vocabulario: (1) selección de la definición que corresponda a la palabra estímulo, (2) selección del ejemplo adecuado para insertar la unidad léxica dada, y (3) escritura de una oración que responda a una pregunta y contenga la palabra estímulo.

2 Los resultados parciales con hablantes de portugués pueden consultarse en Andión y San-Mateo (2018) y en San-Mateo y Chacón (2019), y en San-Mateo y Criado (en prensa) los de los hablantes de francés.

- II. Determinar si el conocimiento que se adquiere con cada actividad es receptivo o productivo.
- III. A la luz de la *HNP* y de la *TAC*, establecer qué criterio predice mejor la eficacia de las actividades de aprendizaje.

Según la *HNP* (Hulstijn y Laufer, 2001: 543-545; Zou, 2017: 70-71), en las tres actividades empleadas (selección de definiciones, selección de ejemplos y escritura de oraciones), la *Necesidad* (N) de identificar o utilizar el significado de la palabra está presente y es moderada (+N), puesto que viene impuesta por un agente externo (motivación extrínseca) y no por el aprendiz, que se limita a seguir instrucciones. Por otro lado, el factor *Búsqueda* (B) está ausente en las tres actividades (-B), pues el sujeto recibe las palabras y las definiciones en el glosario que le facilitamos. En cuanto al tercer factor, la *Evaluación* (E), no existe (-E) en la actividad de selección de definiciones; pero es moderada (+E) en la de selección de ejemplos porque el aprendiz debe reconocer las diferencias entre los contextos dados y seleccionar el adecuado para la palabra estímulo; y es fuerte (++E) en la actividad de escritura de oraciones, ya que implica crear un contexto e incluir la palabra nueva entre otras ya conocidas. Así, la puntuación de las tres actividades, teniendo en cuenta los tres factores que establece la *HNP*, sobre un máximo de 7 puntos, es la siguiente: (1) Selección de definiciones: 1 punto [+N, -B, -E (1 + 0 + 0)], (2) Selección de ejemplos: 2 puntos [+N, -B, +E (1 + 0 + 1)] y (3) Escritura de oraciones: 3 puntos [+N, -B, ++E (1 + 0 + 2)].

Por otro lado, según la *TAC* (Nation y Webb, 2011; Webb y Nation, 2017), las actividades se describen de acuerdo con la respuesta a las 18 preguntas (§ 2) referidas a cinco factores (Tabla 1).

Factores	(1) Selección de definiciones	(2) Selección de ejemplos	(3) Escritura de oraciones
(1) Motivación	2	2	1
(2) Percepción o reconocimiento	2	2	2+1 = 3
(3) Recuperación	3	3	3+1 = 4
(4) Generación	0	2	3
(5) Retención	2	2	2
Puntuación	9/20	11/20	13/20

Tabla 1. Puntuación de las actividades según la *TAC*

De acuerdo con la *Motivación*, las tres poseen un objetivo de aprendizaje de vocabulario claro, pero solo las de selección estimulan el aprendizaje porque suponen “*a challenge to learners*” (Nation y Webb, 2011: 8).

En cuanto a la *Percepción o reconocimiento*, el modo en que se perciben o reconocen las nuevas palabras en la presentación de las actividades se centra la atención en las palabras estímulo y las tres aumentan la conciencia del sujeto de estar aprendiendo nuevo vocabulario, pero de diferente manera: en las actividades de selección de definiciones y de ejemplos solo se realiza en un sentido —el aprendiz debe seleccionar el significado o el contexto correcto entre varias opciones—; en la de la escritura de oraciones, toma contacto con la palabra estímulo en la pregunta planteada y además debe utilizarla en el contexto creado, de ahí que, siguiendo la propuesta de San-Mateo y Criado (en prensa), añadamos un punto más a la tercera actividad (Tabla 1).

Respecto de la *Recuperación*, las tres actividades conllevan restablecer el significado de las palabras, pero no la forma —se trata de una recuperación receptiva, por lo tanto—. Sin embargo, la que conlleva la actividad de selección de ejemplos consiste solo en modificar la palabra para que encaje en el contexto dado, pero en la actividad de escritura es preciso modificarla pues el aprendiz crea el contexto (San-Mateo y Criado, en prensa). Según estos dos niveles de *Recuperación productiva*, adjudicamos un punto más a la tercera actividad (Tabla 1). En las tres el significado de las palabras estímulo está en el glosario facilitado, la misma palabra se recupera dos veces en cada tipo de actividad y de manera no consecutiva.

Por lo que se refiere al factor *Generación*, la selección de ejemplos y la escritura de oraciones implican creación desde un punto de vista productivo, y la última actividad a mayor nivel, pues precisa el dominio de nuevas combinaciones léxicas y de los mecanismos de derivación morfológica.

Finalmente, sobre la *Retención*, las tres actividades aseguran la correcta asociación entre forma y significado, y eluden la interferencia puesto que las palabras estímulo no pertenecen a un mismo campo léxico. Sin embargo, ninguna de las tres implica presencia visual de la palabra en una situación significativa ni tampoco la asociación premeditada de una imagen relacionada con el significado de la palabra.

En definitiva, según la *TAC*, la puntuación de las tres actividades, sobre un máximo de 20³, es esta: (1) Selección de definiciones: 9 puntos; (2) Selección de ejemplos: 11 puntos; y (3) Escritura de oraciones: 13 puntos (Tabla 1).

3 Aunque son 18 preguntas, en la 5 y la 8 se pueden obtener 2 puntos siguiendo a San-Mateo y Criado (en prensa).

Teniendo en cuenta tanto la *HNP* como la *TAC*, las actividades se ordenan de igual forma, si bien esta última permite evaluar un mayor número de aspectos.

Las preguntas de investigación a las que responderemos son las siguientes:

- (1) ¿Cuál es el grado de eficacia para el aprendizaje de vocabulario de las tres actividades propuestas: selección de definiciones, selección de ejemplos y escritura de oraciones?
- (2) ¿En qué medida predicen la eficacia de las actividades estos dos marcos de referencia: la *Hipótesis del Nivel de Participación o Implicación* y la *Técnica de Análisis de Características*?
- (3) ¿Qué tipo de conocimiento: receptivo o productivo fomenta cada actividad?

4 Metodología

4.1 Participantes

Constituyen la muestra 1317 estudiantes de español, 384 hombres (29 %) y 933 mujeres (71 %)⁴, de las siguientes instituciones: Université Sorbonne Nouvelle-Paris 3, Université Paris-Dauphine, École Nationale Supérieure des Mines (París, Francia); Universidade Federal, Universidade Estadual e Instituto Federal (Paráiba, Brasil); City University of New York (CUNY) (Nueva York, Estados Unidos); Università degli Studi di Palermo y Università degli Studi di Messina (Sicilia, Italia); Instituto Cervantes de París, Sao Paulo, Palermo y Nueva York; Universidad Politécnica de Valencia, Universidad Complutense de Madrid, Universidad Popular Miguel Delibes y las Escuelas Oficiales de Idiomas Jesús Maestro (Madrid) y de San Sebastián de los Reyes (España). En total, 1006 sujetos estudian español en contexto heterosiglótico o de LE (un 76 %) y 311, en contexto homosiglótico o de L2 (24 %).

El nivel de español de los participantes es, como mínimo, B1-B2, según el *MCER* (Consejo de Europa, 2002). Ese fue un requisito para participar en el estudio y el 68,4 % (901 sujetos) lo tenía. No obstante, algunos participantes matriculados en cursos de B1-B2 señalaron poseer un nivel Básico (A1-A2):

4 La investigación fue autorizada por el Comité de Bioética de la Universidad Nacional de Educación a Distancia (UNED) al considerarse no invasiva ni perjudicial para los sujetos participantes.

19,4 %, 255 sujetos⁵. El 12,2 % restante (161 sujetos) tenía nivel Superior (C1-C2). Independientemente del nivel que los sujetos indican en el cuestionario de información personal que deben rellenar, se ha verificado con el centro que los cursos en los que están matriculados requieren un nivel B1.

Los participantes poseían 27 lenguas maternas. Los grupos mayoritarios eran hablantes de francés (26,6 %), portugués (23,1 %), italiano (18,1 %) e inglés (16 %); seguidos de los de alemán (4,3 %), chino (4,6 %), polaco (1,1 %) y árabe (1,3 %). Los hablantes del resto de lenguas (búlgaro, checo, coreano, croata, finés, hebreo, holandés, húngaro, japonés, lituano, noruego, persa, ruso, rumano, sueco, tagalo, turco, ucraniano y vietnamita) no llegan al 1 %. En cuanto a la edad, la mayor parte (1141 sujetos, 87 %) tenía entre 15 y 30 años; el 9 % (112), entre 31 y 54, y el 5 % (64), más de 55 años.

El 62 % de los participantes (821) poseía conocimientos de otra lengua además del español, mientras que el 25 % (331) estudiaba dos o más L2. El español era la única L2/LE para el 13 % (165).

4.2 Materiales

Las tres actividades son similares a las de Matanzo (1991), Reyes (1995) y San-Mateo (2005, 2012), y consisten en (1) seleccionar la definición correspondiente a la palabra estímulo (véase el Ejemplo 1), (2) identificar el ejemplo en el que es adecuado insertar la palabra estímulo (Ejemplo 2), y (3) responder a una pregunta planteada con una oración que incluya la palabra estímulo (Ejemplo 3)⁶.

Ejemplo 1. Actividad de selección de definiciones:

¿Qué significa PACTAR? Escribe la palabra al lado de la definición que le corresponda.

1. _____ A propósito, con intención (carácter negativo).
2. pactar Acordar algo entre dos o más personas que están obligadas a cumplirlo.
3. _____ Caer gotas pequeñas de lluvia o de un líquido.
4. _____ Recipiente o vaso en que se conserva y transporta algo.

5 Estos estudiantes, en su mayoría brasileños, demostraron que podían completar las actividades, de ahí que se aceptara su participación.

6 Andión y San-Mateo (2018: 443) confirmaron la idoneidad “para hacer mediciones de eficacia léxica” de los materiales utilizados en el estudio, creados teniendo en cuenta los de Coomber, Ramstad y Sheets (1986) y Matanzo (1991), y del procedimiento empleado.

Ejemplo 2. Actividad de selección de ejemplos:

¿En qué ejemplo usarías la palabra PACTAR? Escribe la palabra en el hueco del ejemplo que le corresponda.

1. Me dijo que no lo hizo _____ ni con mala intención.
2. Creo que debemos ponernos de acuerdo y pactar sobre la cantidad de dinero que pagará cada uno.
3. Hay una mancha en el techo y _____ agua sucia.
4. Dame cualquier _____ para guardar lo que ha sobrado de la paella.

Ejemplo 3. Actividad de escritura de oraciones:

Responde a la siguiente pregunta con una oración completa. Utiliza la palabra indicada en mayúsculas y subráyala. Tu contestación debe demostrar que conoces el significado de dicha palabra.

¿Crees que es adecuado PACTAR con tu pareja una distribución de las tareas domésticas? ¿Por qué?

Sí, me parece muy práctico pactar quién va realizar cada trabajo de la casa porque así evitamos muchos problemas de convivencia.

El pretest o actividad de aprendizaje está compuesto de 12 ítems (uno por palabra estímulo), y hay tres modelos: en cada uno se trabaja con un tipo de actividad (selección de definiciones, de ejemplos o escritura de oraciones). Por otro lado, el test utilizado para verificar la incorporación de las 12 palabras estímulo al léxico mental de los sujetos está dividido en tres secciones de 12 ítems (36 preguntas). En cada sección se practica con uno de los tres tipos de actividad. El formato de los ítems del pretest y el test es el mismo (véanse los ejemplos anteriores).

El glosario está formado por 24 palabras de 6 letras, sin tilde y de 4 categorías gramaticales (*nombre, adjetivo, verbo y adverbio*; 6 de cada). 12 son palabras de baja frecuencia en español y se utilizan como distractores (*adrede, aprisa, dorsal, envase, gotear, hogaño, inicuo/-a, muleta, pactar, pecoso/-a, ubicar y viruta*) y 12 son palabras estímulo (*barmil, catilo/-a, diforo/-a, ecivar, jotone, letaer, nesoal, ronoar, sotiro, talefa, urjale y zienga*), pseudopalabras creadas siguiendo los patrones fonológicos y de formación de palabras, y la frecuencia de las estructuras silábicas del español⁷.

7 Véase una descripción detallada de la selección de las palabras-distractores, la creación de las palabras estímulo y de la asignación del significado en Andión y San-Mateo (2018: 442–443), San-Mateo y Chacón (2019: 5–7) y San-Mateo y Criado (en prensa).

Finalmente, la información personal y sociolingüística de los participantes (edad, sexo, dominio de la L2 y de otras lenguas, etc.) se recoge en un cuestionario que completan entre la fase de aprendizaje y el test.

4.3 Procedimiento

La recopilación de datos se realizó en las instituciones de Brasil, España, Estados Unidos, Francia e Italia antes mencionadas, siguiendo el procedimiento de Coomber, Ramstad y Sheets (1986), en una sesión de 45 minutos. En primer lugar, los participantes firman el consentimiento de participación. Después, el investigador reparte el glosario de palabras y lee en voz alta las 24 palabras y sus definiciones. De manera aleatoria, asigna a cada estudiante uno de los tres tipos de actividad de aprendizaje. En total, 430 estudiantes practicaron con la actividad de selección de definiciones (grupo 1), 438 con la de ejemplos (grupo 2) y 449 con la de escritura de oraciones (grupo 3).

El tiempo para completar cada uno de los 12 ítems del pretest es de 45 segundos. Pasado ese tiempo, el investigador lee la definición correcta de la palabra con la que se está trabajando (retroalimentación) para que el sujeto compruebe su respuesta. En esta fase, la actividad se repite dos veces, pero solo en la primera está permitida la consulta del glosario. Así la exposición a las palabras estímulo es doble y el esfuerzo requerido en la segunda serie de actividades es mayor al no poder usar el glosario. Para evitar el efecto de *recencia* que podría favorecer recordar la última palabra de la secuencia (Gavett y Horwitz, 2012), antes de completar el test los participantes responden al cuestionario de información personal durante 5-10 minutos.

El test final que mide la retención de las palabras estímulo es igual para todos los participantes, que trabajan a su ritmo durante 15 minutos, sin acceso al glosario y sin retroalimentación. Deben responder a 12 ítems de cada tipo de actividad: en la sección 1 trabajan con la selección de definiciones; en la 2, con la de ejemplos; y en la 3, con la escritura de oraciones. La puntuación máxima en cada sección del test es de 12 puntos: se asigna un punto por cada respuesta correcta y cero por las incorrectas o en blanco. Este sistema de puntuación es el seguido en estudios similares (Agustín, 2009; Folse, 2006; Hulstijn y Laufer, 2001; San-Mateo, 2012 y Webb, 2005; entre otros).

5 Resultados

5.1 Estadística descriptiva

La puntuación media del test, es decir, el número medio de palabras que los sujetos (1317) son capaces de recordar es 9,11 sobre 12 puntos, que equivale a un 75,92 %. Esta elevada cifra prueba la eficacia de las tres actividades de aprendizaje para ayudar a los estudiantes a ampliar el vocabulario en la L2. La pequeña desviación estándar (2,69) implica que la habilidad de los sujetos para completar las actividades es homogénea.

5.2 Resultados de la variable actividad de aprendizaje

Para responder a la primera pregunta de investigación, analizamos los resultados de cada una de las actividades de aprendizaje (variable independiente). El número de palabras que los sujetos son capaces de recordar —es decir, los aciertos en el test— es la variable dependiente. El grupo entrenado con la selección de definiciones ha obtenido una puntuación de 7,96 (66,37 %); el que ha trabajado con la selección de ejemplos, de 9,06 (75,53 %); y el grupo que ha practicado con la escritura de oraciones, de 10,25 (85,43 %), la puntuación más elevada (Tabla 2). Por lo tanto, esta última actividad es la más eficaz porque ha permitido a los sujetos recordar una mayor cantidad de palabras en el test.

Actividad	N	Media*	Desviación estándar	Error estándar
(1) Selección de definiciones	430	7,964	2,864	0,138
(2) Selección de ejemplos	438	9,064	2,708	0,129
(3) Escritura de oraciones	449	10,252	1,916	0,090

Nota: *Puntuación máxima = 12

Tabla 2. Media y desviación estándar de la puntuación del test según la actividad de aprendizaje

Con el fin de comprobar si es significativa la diferencia entre las puntuaciones obtenidas por cada grupo establecido según la actividad de aprendizaje, hemos llevado a cabo un análisis de varianza (ANOVA)⁸. Este indica que la diferencia

8 Este análisis estudia las diferencias entre las medias de los grupos. El resultado es estadísticamente significativo cuando resulta improbable que las diferencias hayan ocurrido por casualidad. El nivel de significatividad o umbral de probabilidad (valor de *p*) es 0,05; cuando este valor es igual o menor que 0,05, las diferencias entre las medias se consideran estadísticamente significativas. Para comprobar entre qué grupos las diferencias son significativas, se

es estadísticamente significativa con un nivel de confianza del 95 %: $F(2, 1314) = 90,322$, $p = 0,000$. Dado que no hay homogeneidad de varianzas, se aplica la prueba de Games-Howell e indica que la diferencia entre las medias obtenidas en el test por cada grupo es significativa en un nivel de confianza del 95 %: p es 0,000 en los tres casos: entre el grupo que ha practicado con la selección de definiciones y los otros dos; y entre el grupo que ha trabajado con la selección de ejemplos y el que se ha entrenado con la escritura de oraciones.

Estos datos confirman que la actividad de escritura de oraciones conlleva un resultado mucho mejor que la selección de ejemplos y la de definiciones; y que la actividad de selección de ejemplos, un resultado mejor que la de elegir la definición. Por lo tanto, existen diferencias en el aprendizaje de las nuevas palabras si se practica con cada una de las actividades.

5.3 Resultados del test por secciones

En primer lugar, calculamos el promedio de aciertos de todos los participantes en cada una de las tres secciones del test y observamos que es más alto en la sección en la que deben seleccionar la definición de la palabra estímulo (10,11 aciertos sobre 12: 84,23 %); en segundo lugar, en la que hay que completar los ejemplos (8,96 aciertos de media: 74,65 %) y, por último, solo son capaces de utilizar correctamente 8,26 palabras (68,87 %) en la que deben responder a la pregunta planteada. La diferencia entre estas puntuaciones es estadísticamente significativa según el modelo lineal general ($p = 0,000$), lo cual quiere decir que las tres actividades empleadas para el aprendizaje de las palabras capacitan al sujeto sobre todo para realizar tareas de reconocimiento (selección de definiciones y de ejemplos) en las que es preciso solo un conocimiento receptivo del vocabulario, y bastante menos para las de producción.

En segundo lugar, para responder a la tercera pregunta de investigación sobre el tipo de conocimiento (receptivo o productivo) de las palabras que fomenta cada actividad, nos centraremos en las puntuaciones de cada sección del test teniendo en cuenta la actividad de aprendizaje realizada. Así analizamos si las actividades receptivas (selección de definiciones y de ejemplos) entrenan a los estudiantes para llevar a cabo también actividades productivas;

llevan a cabo comparaciones múltiples *post hoc* y el valor de p también es 0,05. Dependiendo de si las varianzas son iguales o desiguales (según el estadístico de Levene), se elige el tipo de prueba *post hoc* (Bonferroni, HSD Tukey, Scheffé, Tamhane, Games-Howell, entre otras). Los cálculos estadísticos se han realizado con el programa IBM SPSS Statistics para Windows (IBM Corp., 2016).

y si la productiva (escritura de oraciones) los entrena para completar las receptivas.

Como muestra la Tabla 3, el grupo que ha trabajado con la selección de definiciones es el que menor puntuación alcanza en todas las secciones del test y es especialmente baja en la tercera, en la que debe emplear de manera productiva las palabras aprendidas: 6,41 palabras, es decir, apenas un 54 % del total. Por el contrario, la puntuación de los sujetos que practicaron con la actividad de escritura es siempre más alta que la de los otros dos en las tres secciones del test. Es decir, habiendo trabajado con una actividad productiva, los resultados son mejores tanto en la productiva como en las receptivas.

Actividad de aprendizaje	Test		
	Sección 1: Selección de definiciones Media (DE)	Sección 2: Selección de ejemplos Media (DE)	Sección 3: Escritura de oraciones Media (DE)
(1) Selección de definiciones ($n = 430$)	9,237 (3,243)	8,249 (3,174)	6,407 (3,915)
(2) Selección de ejemplos ($n = 438$)	10,085 (2,691)	8,884 (3,254)	8,224 (3,605)
(3) Escritura de oraciones ($n = 449$)	10,964 (1,901)	9,708 (2,785)	10,082 (2,549)

Tabla 3. Media y desviación estándar de la puntuación en cada sección del test según la actividad de aprendizaje

El análisis de varianza indica que la diferencia entre las puntuaciones obtenidas en cada sección del test teniendo en cuenta la actividad de aprendizaje es significativa en los tres casos, puesto que el valor de p menor de 0,05: en la sección 1 (selección de definiciones), $F(2, 1314) = 46,338$, $p = 0,000$; en la 2 (selección de ejemplos), $F(2, 1314) = 24,928$, $p = 0,000$; y en la 3 (escritura de oraciones), $F(2, 1314) = 128,571$, $p = 0,000$. Para saber entre qué grupos se da la diferencia debemos realizar las comparaciones múltiples *post hoc*. El test de Games-Howell, en concreto, porque las varianzas no son homogéneas, muestra que en las tres secciones el efecto de las tres actividades de aprendizaje es significativo con un nivel de confianza del 95 % (en las secciones 1, 2 y 3: $p = 0,000$ en los tres).

En resumen, en las tres secciones del test la actividad de escritura ha sido la que más ha facilitado la retención de nuevas palabras (Figura 1). Y las diferencias entre las tres actividades de aprendizaje son significativas cuando, en el test, los sujetos deben identificar la definición correcta de la palabra y el ejemplo donde se puede insertar, y escribir una oración que incluya la palabra estímulo.

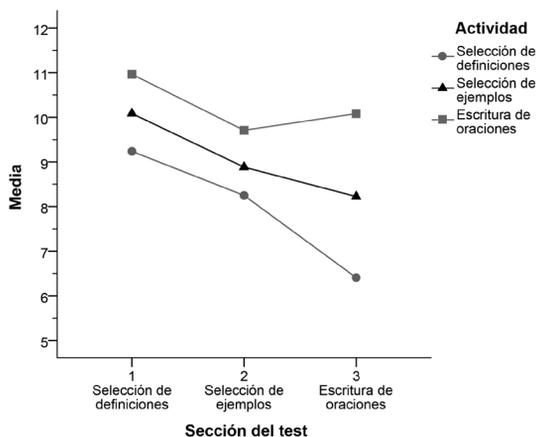


Figura 1. Puntuación media en cada sección del test según la actividad de aprendizaje

6 Discusión

En relación con la primera pregunta de investigación: ¿Cuál es el grado de eficacia para el aprendizaje de vocabulario de las tres actividades propuestas: selección de definiciones, de ejemplos y escritura de oraciones?, los resultados indican que, aunque las tres actividades favorecen el aprendizaje de vocabulario en la L2 (la media de aciertos ha sido de 9,11 sobre 12; un 75,92 %), se pueden establecer grados de efectividad: la más eficaz ha sido la escritura de oraciones, seguida de la selección de ejemplos y de la selección de definiciones.

Esta escala es precisamente la que predice la *HNP* y la *TAC* (pregunta n.º 2). Según la primera, la actividad que requiere un mayor nivel de participación o esfuerzo cognitivo –en nuestro caso, la escritura de oraciones (+N, –B, ++E)– es la más eficaz para aprender nuevas palabras (porcentaje de aciertos: 85,43 %). Por el contrario, la selección de definiciones es la actividad que menor nivel de participación implica (+N, –B, –E) y es, asimismo, la menos eficaz (porcentaje de aciertos: 66,37 %). La selección de ejemplos se sitúa entre las otras dos actividades en cuanto a nivel de participación (+N, –B, +E) y en los aciertos en el test (75,53 %).

Según la *TAC*, la actividad de selección de definiciones es también la que menor puntuación obtiene en la escala y la menos eficaz en nuestro estudio. Por otro lado, si aplicamos las sugerencias de San-Mateo y Criado (en prensa) a los criterios propuestos por Nation y Webb (2011), en concreto al 2 (*Percepción*

reconocimiento) –en la pregunta 5– y al 3 (*Generación*) –en la pregunta 8 –, la puntuación de la actividad de escritura de oraciones es mayor que la que obtiene la selección de ejemplos (13 y 11, respectivamente). Esta escala de gradación está en consonancia con los resultados obtenidos en el estudio. Sin embargo, si no se tienen en cuenta esas modificaciones, la puntuación de las dos actividades mencionadas es la misma (11 puntos) y la técnica de análisis, por lo tanto, no las diferenciaría.

En este sentido, ambos criterios, la *HNP* y la *TAC*, predicen correctamente el nivel de eficacia de las actividades con las que hemos trabajado aquí y no se pueden establecer diferencias entre ellos.

Respecto de la tercera pregunta de investigación: ¿Qué tipo de conocimiento: receptivo o productivo fomenta cada actividad?, hemos comprobado que la actividad para la que los estudiantes demuestran estar más capacitados es la selección de definiciones: les resulta más fácil y han obtenido mayor puntuación en ella (10,11 aciertos sobre 12: 84,23 %). En el extremo opuesto está la actividad de escritura, de mayor dificultad y en la que, de media, aciertan 8,26 palabras (68,87 %). Según la clasificación de Nation (2001) de los tipos de conocimiento y los aspectos que engloban, la selección de definiciones –y también la de ejemplos– requiere solamente conocimiento receptivo del significado de la palabra; en cambio, la escritura de oraciones, previa comprensión de la pregunta que se debe responder, requiere conocimiento receptivo y productivo del significado y del uso de la palabra.

Como actividad de entrenamiento, la selección de definiciones ha sido la menos eficaz si tenemos en cuenta que los aprendices que han practicado con ella son los que peor puntuación obtuvieron tanto en las actividades que necesitan conocimiento receptivo –selección de definiciones y de ejemplos– como en la que exige conocimiento productivo de la palabra –escritura de oraciones–, en la cual la puntuación fue la más baja de todas (6,41 sobre 12). En cambio, el trabajo con la actividad productiva de escritura ha hecho posible que los sujetos realizaran con mayor éxito todas las actividades, las de conocimiento receptivo y productivo: el grupo 3 ha obtenido las mejores puntuaciones en las tres secciones del test, y son significativamente más altas que las de los otros dos grupos.

Por lo tanto, aunque el entrenamiento receptivo posibilita que el aprendiz sea capaz de llevar a cabo actividades productivas, con el entrenamiento productivo consigue no solo reconocer más palabras sino hacer tareas más complejas con ellas, como redactar una respuesta que la incluya.

7 Conclusión

Los resultados del presente estudio están en la línea de otros anteriores que avalan la eficacia de las actividades de escritura para incorporar nuevas palabras al lexicón mental del estudiante de L2 (cf. Alcaraz y Almela, 2013; Folse, 2006), y de la capacidad de predicción de la *HNP*, que establece que las actividades que precisan mayor nivel de participación o esfuerzo cognitivo por parte del aprendiz son las más eficaces (Agustín, 2009; Alavinia y Rahimi, 2019; Browne, 2003; Hulstijn y Laufer, 2001; Keating, 2008; Kim, 2008; Kondo, 2007; Laufer, 2003; Pichette, De Serres y Lafontaine, 2012; San-Mateo, 2012; Tahmasbi y Farvardin, 2017; y Webb, 2005; entre otros). También se confirma la eficacia de predicción de la *TAC*, pues, al incorporar los dos niveles de *Percepción o reconocimiento* y de la *Recuperación productiva* (San-Mateo y Criado, en prensa), se distinguen las actividades de escritura y de selección de ejemplos. No podemos afirmar que esta última escala sea “*a significantly stronger predictor of lexical gains*”, como en Hu y Nassaji (2016: 37), puesto que, en nuestro caso, ambas muestran el mismo poder explicativo.

En esta investigación, la actividad de escritura, de mayor puntuación en las escalas del *Nivel de Participación o Implicación* y del *Análisis de Características*, es con la que mejores resultados consiguen los estudiantes en actividades productivas y receptivas; en cambio, con la actividad receptiva de selección de definiciones, la de menor puntuación en ambas escalas, los resultados son siempre peores, especialmente en la actividad productiva. Por lo tanto, a la hora de seleccionar actividades es importante que el docente tenga en cuenta el tipo y grado de conocimiento perseguido, y la tarea elegida.

El hecho de que la muestra de esta investigación sea tan amplia y variada: 1317 informantes con nivel de español mínimo B1-B2, procedentes de países e instituciones diversas, hablantes de 27 lenguas maternas, etc., implica que los resultados obtenidos son una aportación importante en el conocimiento del proceso de aprendizaje del vocabulario en L2 y que pueden extrapolarse a numerosos contextos de aprendizaje.

Agradecimientos

Agradecemos la inestimable colaboración de los equipos docentes y de los alumnos de los centros participantes, y la de Cecilia Criado y M.^a José Labrador, del Grupo de investigación INVOLEX.

Bibliografía

- Agustín, M. P. (2009): «The effect of reading only, reading and comprehension, and sentence writing in lexical learning in a foreign language: Some preliminary results». *RESLA*, 22, 9-33.
- Alavina, P., Rahimi, H. (2019): «Task types effects and task involvement load on vocabulary learning of EFL learners». *International Journal of Instruction*, 12(1), 1501-1516.
- Alcaraz, G., Almela, A. (2013): «The Involvement Load Hypothesis: Its effect on vocabulary learning in Primary Education». *RESLA*, 26, 11-24.
- Andión, M.ª A., San-Mateo, A. (2018): «Tres actividades de adquisición léxica. Resultados en aprendices brasileños de español como lengua extranjera». *Calidoscópio*, 16(3), 437-449.
- Browne, C. (2003): *Vocabulary Acquisition through Reading, Writing and Tasks: A Comparison*. Tesis doctoral. Japan: Temple University Japan.
- Chaharlang, N., Farvardin, M. T. (2018): «Predictive power of Involvement Load Hypothesis and Technique Feature Analysis across L2 vocabulary learning tasks». *International Journal of Foreign Language Teaching and Research*, 6(24), 127-141.
- Consejo de Europa (2002): *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza y evaluación (MCER)*. Madrid: Secretaría General Técnica del MECD-Subdirección General de Información y Publicaciones, y Grupo ANAYA, S. A.
- Coomber, J. E., Ramstad, D. A., Sheets, D. A. R. (1986): «Elaboration in vocabulary learning: A comparison of three rehearsal methods». *Research in the Teaching of English*, 20(3), 281-293.
- Folse, K. S. (2006): «The effect of type of written exercise on L2 vocabulary retention». *TESOL Quarterly*, 40(2), 273-293.
- Gavett, B. E., Horwitz, J. E. (2012): «Immediate list recall as a measure of short-term episodic memory: Insights from the serial position effect and item response theory». *Archives of Clinical Neuropsychology*, 27(2), 125-135.
- Gohar, M. J., Rahmanian, M., Soleimani, H. (2018): «Technique Feature Analysis or Involvement Load Hypothesis: Estimating their predictive power in vocabulary learning». *Journal of Psycholinguistic Research*, 47(4), 859-869.

- Hu, H. M., Nassaji, H. (2016): «Effective vocabulary learning tasks: Involvement Load Hypothesis versus Technique Feature Analysis». *System*, 56, 28-39.
- Hulstijn, J. H., Laufer, B. (2001): «Some empirical evidence for the Involvement Load Hypothesis in vocabulary acquisition». *Language Learning*, 51(3), 539-558.
- IBM Corp. (2016): *IBM SPSS Statistics para Windows, Versión 24.0*. Nueva York: IBM Corp.
- Keating, G. D. (2008): «Task effectiveness and word learning in a second language: The Involvement Load Hypothesis on trial». *Language Teaching Research*, 12(3), 365-386.
- Khoshsima, H., Eskandari, Z. (2017): «Task effectiveness predictors: Technique Feature Analysis vs. Involvement Load Hypothesis». *Iranian Journal of English for Academic Purposes*, 6(2), 50-69.
- Kim, Y. (2008): «The role of task-induced involvement and learner proficiency in SL vocabulary acquisition». *Language Learning*, 58(2), 285-325.
- Kondo, H. (2007): «The effects of semantic elaboration on SL vocabulary learning». *Research Journal of Jin-Ai University*, 6, 71-78.
- Laufer, B. (2003): «Vocabulary acquisition in a second language: Do learners really acquire most vocabulary by reading? Some empirical evidence». *The Canadian Modern Language Review*, 59(4), 567-587.
- Laufer, B. (2017): «From word parts to full texts: Searching for effective methods of vocabulary learning». *Language Teaching Research*, 21(1), 5-11.
- Matanzo, G. (1991): *Vocabulario y enseñanza: Estudio de la relación existente entre los métodos empleados y la incorporación del léxico nuevo a la competencia lingüística de estudiantes universitarios puertorriqueños*. Tesis doctoral. San Juan: Universidad de Puerto Rico.
- Nation, P. (2001): *Learning vocabulary in another language*. Cambridge: CUP.
- Nation, P., Webb, S. (2011): *Researching and analyzing vocabulary*. Boston: Heinle.
- Pichette, F., Serres, L. de, Lafontaine, M. (2012): «Sentence reading and writing for second language vocabulary acquisition». *Applied Linguistics*, 33(1), 66-82.
- Reyes, M.^a J. (1995): *Enriquecimiento de la competencia léxica: análisis estadístico*. Tesis doctoral. Las Palmas: Universidad de Las Palmas de Gran Canaria.

- San-Mateo, A. (2005): *Aprendizaje de léxico en español como segunda lengua. Investigación sobre tres métodos*. Trabajo Final de Máster. Madrid: UNED.
- San-Mateo, A. (2012): *Aprendizaje de léxico en español como segunda lengua/lengua extranjera: investigación sobre la efectividad de tres tipos de actividades para aprender vocabulario*. Tesis doctoral. Madrid: UNED.
- San-Mateo, A., Chacón, C. (2019): «Learning word class in a second language through vocabulary learning activities: definition-choosing, gap-filling, and sentence-writing». *Journal of Spanish Language Teaching*, 6(1), 49-63.
- San-Mateo, A., Criado, C. (en prensa): «Receptive and productive vocabulary acquisition: effectiveness of three types of tasks. Results from French students of Spanish». *Onomazéin*.
- Tahmasbi, M., Farvardin, M. T. (2017): «Probing the effects of task types on EFL learners' receptive and productive vocabulary knowledge: The case of Involvement Load Hypothesis». *SAGE Open*, 7(3), 1-10.
- Webb, S. (2005): «Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge». *Studies in Second Language Acquisition*, 27(1), 33-52.
- Webb, S., Nation, P. (2017): *How vocabulary is learned*. Oxford: OUP.
- Zou, D. (2017): «Vocabulary acquisition through cloze exercises, sentence-writing and composition writing: Extending the evaluation component of the Involvement Load Hypothesis». *Language Teaching Research*, 21(1), 54-75.
- Zou, D., Wang, F. L., Kwan, R., Xie, H. (2018): «Investigating the effectiveness of vocabulary learning tasks from the perspective of the Technique Feature Analysis: The effects of pictorial annotations». En: Simon K. S. Cheung, Jeanne Lam, Kam Cheong Li, Oliver Au, Will W. K. Ma, Wai Shing Ho (eds.), *Technology in Education. Innovative Solutions and Practices. ICTE 2018. Communications in Computer and Information Science*. New York: Springer, 3-15.

Alicia San-Mateo Valdehíta and

M.^a Antonieta Andi3n Herrero

National Distance Education University (UNED)

Efficacy of three vocabulary-learning activities in Spanish FL/L2: selection of definitions, selection of examples and the writing of sentences

Keywords: Spanish as a Second Language/Foreign Language, vocabulary learning, learning activities, Involvement Load Hypothesis, Technique Feature Analysis.

This article presents the overall results of a research project on vocabulary learning in Spanish, carried out with 1,317 participants, from 27 mother tongues, in the context of it being foreign language and second language. The aim is to determine (1) which of these three vocabulary-learning activities is more effective: definition-choosing, gap-filling or sentence-writing; (2) the predictive ability of two scales with regard to effective vocabulary-learning activities: *Involvement Load Hypothesis* (Hulstijn and Laufer, 2001) and *Technique Feature Analysis* (Nation and Webb, 2011); and (3) what kind of knowledge, receptive or productive, learners acquired with each activity. The results showed that subjects who performed sentence-writing, the productive activity, which required more cognitive effort according to the two scales, remembered more words and were able to complete receptive and productive activities significantly better than the other groups. In contrast, those who performed definition-choosing, which entailed only receptive knowledge of the words and ranked less on both scales, remembered fewer words, and had worse results with regard to productive activity.

Alicia San-Mateo Valdehíta in
M.^a Antonieta Andi3n Herrero

Nacionalna univerza za študij na daljavo (UNED)

U3inkovitost treh aktivnosti za u3enje besediš3a v španš3ini kot tujem/drugem jeziku: izbira prave razlage, izbira primera rabe in pisanje povedi

Ključne besede: španš3ina kot drugi/tuji jezik, u3enje besediš3a, u3ne aktivnosti, *Involvement Load Hypothesis*, *Technique Feature Analysis*

V prispevku se predstavi splošne rezultate raziskave o u3enju besediš3a v španš3ini kot tujem in drugem jeziku, v kateri je sodelovalo 1317 udeleženecv s 27 maternimi jeziki. Namen je dolo3iti, (1) katera od teh treh aktivnosti za u3enje besediš3a (izbira prave razlage besede, izbira primera, v katerem se lahko uporabi dana beseda, in pisanje povedi z dano besedo) je najbolj u3inkovita, (2) kako natan3no lahko lestvici *Involvement Load Hypothesis* (Hulstijn in Laufer, 2001) in *Technique Feature Analysis* (Nation in Webb, 2011) napoveva u3inkovitost vseh treh aktivnosti in (3) kakšno znanje, receptivno ali produktivno, se pridobi z vsako od aktivnosti. Rezultati kažejo, da so si udeleženeci, ki so besediš3e pridobivali ali vadili s pisanjem, torej s produktivno aktivnostjo, ki po obeh lestvicah zahteva najve3 kognitivnega napora, zapomnili najve3 besed in da so aktivnosti iz receptivnega in produktivnega znanja reševali veliko bolje od ostalih. Tisti, ki pa so vadili z izbiro prave razlage besede, ki predvideva le receptivno znanje besede in na obeh lestvicah dosega najmanj to3k, pa so si zapomnili najmanj besed in imeli pri produktivnih aktivnostih najslabše rezultate.