

VALORACIÓN DEL APRENDIZAJE EN LA TRADUCCIÓN

M^a JESÚS SÁNCHEZ MANZANO
Universidad de Salamanca

RESUMEN. *El propósito de esta investigación era evaluar los cambios producidos en la traducción en función del aprendizaje. Se utilizaron dos grupos de sujetos de la Universidad de Salamanca. El primer grupo (Pre/Post-estudio) participó en las dos fases del experimento, pero no se le informó de la tarea que se iba a ejecutar en la segunda fase. Los estudiantes del grupo Post-estudio participaron sólo en la segunda fase. En cada fase, los estudiantes realizaron una tarea de traducción y otra de juicios de relación. Se recogieron juicios de relación de un grupo experto, y también se analizaron con el algoritmo Pathfinder. Sólo se obtuvo evidencia de aprendizaje en las valoraciones originales y en la tarea de traducción. Dado que la similitud entre las representaciones estructurales de los expertos y las de los estudiantes no proporcionó evidencia de aprendizaje, probablemente se trate de un aprendizaje superficial.*

PALABRAS CLAVE. *Pathfinder, traducción, evaluación, aprendizaje, relación, similitud, valoración, puntuación.*

ABSTRACT. *The aim of this investigation was to assess changes in the translation as a function of learning. Two groups of subjects from the University of Salamanca were run. The first group, the Prestudy-Poststudy group, participated in the two phases of the experiment; but they were not informed about the task to be performed in the second phase. Students in the Poststudy group participated only in the second phase. In each phase, students performed a translation and a rating task. Relatedness ratings were collected from a group of experts, and were also analyzed with the Pathfinder procedure. Evidence of learning was only got through raw ratings and the scores in the translation task. However, since the similarity between the experts' structural representations and those of the students did not provide evidence of learning, probably we are dealing with a superficial learning.*

KEYWORDS. *Pathfinder, translation, assess, learning, relatedness, similarity, rating, score.*

1. INTRODUCCIÓN

La presente investigación surge de la necesidad de explorar una nueva metodología en la evaluación y en el aprendizaje de idiomas. Se piensa que si se encuentra un método fiable para la evaluación del aprendizaje y del conocimiento,

éste servirá como *feedback* para llegar a conocer las necesidades reales de los estudiantes en cuanto a la materia a aprender. Para ello, se parte de trabajos que evalúan distintas formas y métodos de medir las relaciones estructurales de los estudiantes (Preece 1976; Diekhoff 1983; Gonzalvo, Cañas y Bajo 1994; Johnson, Goldsmith y Teague 1995, etc.). Entre estos autores existe la idea de que el aprendizaje conlleva incorporar nueva información al conocimiento que ya se tiene, lo cual implica nuevas interrelaciones y un cambio en la organización. Dentro de esta concepción, los hallazgos de Gonzalvo *et al.* (1994) y los de Johnson *et al.* (1995), son decisivos al emprender esta investigación. Ellos encuentran que la ejecución del estudiante se predice mejor con las soluciones halladas con el algoritmo Pathfinder (se describe en la sección de metodología y más extensamente en Schvaneveldt, Durso, Goldsmith, Breen, Cooke, Tucker y De Maio 1985) que con los datos originales o con otras medidas, como las que se obtienen con los exámenes. A la luz de esas investigaciones, tres son las medidas o variables dependientes que aquí se utilizan para intentar captar cambios cuantitativos por el estudio. Estas son: las puntuaciones obtenidas en una prueba de traducción, la medida de proximidad conseguida a través de las valoraciones dadas por los estudiantes en una prueba en la que tenían que valorar la relación existente entre los términos que se les presentaban en una pantalla de ordenador y, finalmente, la medida de similitud. Es decir, la similitud en cuanto al número de enlaces comunes entre los nodos en diferentes redes. Esta medida se halló sometiendo al Pathfinder las valoraciones originales proporcionadas por los estudiantes. Se parte de la idea de que cuando uno se hace más familiar con un determinado vocabulario en una lengua extranjera éste se debería de reorganizar. Por ello se tendría que notar al manipular la variable independiente, en este caso el estudio, un cambio en su organización. Cambio producido por el aprendizaje, el cual se tendría que reflejar en una diferencia cuantitativa.

2. METODOLOGÍA

2.1. Método. Pathfinder

El Pathfinder es el método que se va a utilizar a lo largo de esta investigación. Es un método indirecto de elicitación del conocimiento en el que se utilizan datos de proximidad. Por lo general se le da al sujeto la tarea de relacionar numéricamente parejas de conceptos, en base a las relaciones conceptuales que éstas le sugieren. Normalmente se utiliza una escala que va del 1 al 9, el 1 expresa una escasa relación o ninguna entre la pareja de conceptos presentada, mientras que el 9 expresa la máxima relación entre ellos. El sujeto valora todas las posibles combinaciones de términos, tomados de dos en dos [$n * (n - 1) / 2$, n = número de conceptos], con la primera impresión que le producen. Para generar una red PFNET hay que tener en cuenta los

dos parámetros usados por el Pathfinder: r y q . El parámetro r determina cómo hallar el valor (*weight*) entre una secuencia de nodos (path) a partir de los valores de los enlaces en la secuencia. El parámetro q limita el número de enlaces que puede haber entre una secuencia de nodos. Dependiendo de los valores que se le asignen a estos dos parámetros se puede, con el mismo grupo de datos de proximidad, formar una gran variedad de redes. Los parámetros más usados para generar las redes menos densas, en las investigaciones que tratan con el algoritmo Pathfinder, son con $r =$ infinito, y $q = n - 1$. Con los parámetros establecidos para r y q el Pathfinder produce una red a partir de los datos de proximidad, en la que cada nodo se conecta a otro mediante un puntero o enlace. La distancia entre ambos nodos viene dada por el valor o fuerza asignado a cada enlace, representando los números mayores una distancia psicológica mayor entre los conceptos. Para que el enlace permanezca en la red es condición indispensable, del algoritmo Pathfinder, que constituya la distancia mínima entre los dos nodos.

En la investigación llevada a cabo, se aplicó el Pathfinder con el fin de obtener redes a partir de los datos dados por los sujetos. Los parámetros usados para generar las redes Pathfinder fueron con $r =$ infinito, y $q = n - 1$; en ellas los conceptos aparecían como nodos y sus relaciones, como ya se ha mencionado, venían representadas por enlaces. La fuerza entre los nodos venía marcada por el valor que se le había asignado al enlace, reflejando la distancia entre los nodos a lo largo del enlace.

2.2. Sujetos

Se parte de los presupuestos teóricos de que se pueden detectar cambios en la traducción y en la organización del conocimiento al pasar de una fase a otra de aprendizaje. Por ello se utilizaron dos grupos de sujetos tomados de la misma clase; su pertenencia a un grupo o a otro se hizo de forma aleatoria. Los dos grupos lo componían estudiantes de 1º de Psicología de la Universidad de Salamanca, matriculados en la asignatura de libre disposición *Inglés aplicado a la Psicología*. Para uno se tomaron dos medidas: antes y después del estudio (Pre y Post I); para el otro sólo se tomó una medida una vez estudiado el material asignado (Post II). En condiciones normales se contaría con una medida Pre y Post (de un grupo que no hubiera recibido el tratamiento de la variable independiente: en este caso el estudio), para poder comparar los resultados con el grupo que sí lo había recibido. Sin embargo en este experimento no se podía llevar a cabo debido a que la variable independiente se tenía que aplicar a toda la clase, y no se contaba con ningún otro grupo de sujetos del mismo nivel lingüístico que éste. La medida sólo Post (II), ya que el experimento no se pudo llevar a cabo con un diseño Pretest-Posttest, se tomó exclusivamente con el fin de poder discriminar si existía un efecto de aprendizaje debido a haber realizado la misma tarea en otra ocasión. En el grupo Pre/Post-estudio participaron 33 sujetos, pero se eliminaron 4 debido a que no participaron en la fase final. En cuanto al grupo

Post-estudio participaron 25 sujetos. Se puede decir que los sujetos eran bastante homogéneos en cuanto al conocimiento de la lengua inglesa, ya que en la prueba de nivel realizada al comienzo de curso no se observaron diferencias significativas $t(52) = 0,201, p = 0,8412$.

A los estudiantes voluntarios que intervinieron en la fase Pre-estudio se les ofreció crédito (*extra credit*) por su colaboración. Sin embargo, los sujetos de las fases Post-estudio (I y II) participaron sin saber que era parte de un experimento. Ya que a todos los sujetos, incluso a los que habían participado en la primera prueba un par de meses antes, se les realizó la prueba Post-estudio como parte de un examen parcial. No se les había informado de que el material que aparecería en una sección de este examen sería el mismo que el presentado en la fase Pre-estudio. Así se evitaba que concentraran su estudio en este material, y se obtuvieran datos distorsionados. También se contó con 3 expertos en la enseñanza de la lengua inglesa, como referente ideal con el que comparar los resultados de los estudiantes. Aunque parece un número escaso, se cuenta con investigaciones realizadas con una cantidad reducida de expertos que han proporcionado una estructura referente estable susceptible de ser comparada con la de los estudiantes (Goldsmith, Johnson y Acton 1991; Gonzalvo *et al.* 1994). Se esperaba que la manera de traducir los términos y la estructura conceptual de los estudiantes cambiarían de la fase Pre a la Post I, haciéndose así más similares a las del panel experto.

2.3. Materiales

Dado que los estudiantes iban a ser examinados, dentro de la asignatura *Inglés aplicado a la Psicología*, del contenido y traducción del artículo de Ericsson y Charness: *Expert Performance. Its Structure and Acquisition* (1994) se seleccionó una serie de términos procedentes de este artículo. Para evitar el sesgo de la persona encargada de la elaboración de esta prueba, se utilizó el programa Hyperbase 2.3 para el ordenador Macintosh. Este programa proporcionó, de acuerdo con la longitud de dicho artículo, todas las palabras que aparecían con una frecuencia de 8 o más. Con lo cual se eligieron un total de 26 términos, los cuales presentan entre sí una gran relación (véase la lista de términos en el apéndice final). Se desecharon un buen número de términos relevantes dentro del campo semántico elegido: *Expert Performance*, sobre todo los más similares al español. Dado que la tarea tenía que ver con la traducción, se pensó que en el caso de incluir estos conceptos no se captaría de manera adecuada si se había producido o no el aprendizaje de una fase a otra. Por esta razón palabras como: *memory, experience, structure, etc.*, a pesar de aparecer con bastante frecuencia, no se utilizaron en esta investigación. También se eliminaron los adjetivos, incluso algunos de gran frecuencia y muy relacionados con este campo semántico, como: *successful, outstanding, relevant, etc.* Se tomó esta decisión debido a que al ir antepuesto a los nombres, si aparecían solos daría la impresión de que

estaban incompletos, y si no era así probablemente cada sujeto daría una traducción en función de la frecuencia del nombre que normalmente acompañe a dichos términos. Pudiendo ser, que ese nombre le influya de tal modo que la traducción se aleje en gran medida del campo semántico con el que estamos tratando. Por fin se terminó seleccionando, entre nombres y verbos, un total de 26 palabras. Estas formas podían aparecer en el plural o en el singular, e incluso si hablamos de verbos también se eligió alguna forma verbal en tercera persona del singular. Se hizo así, dependiendo de cual de las formas presentaba una mayor frecuencia.

2.4. *Procedimiento*

2.4.1. *Traducción y definición de los conceptos*

A los sujetos se les pasó un cuadernillo con los 26 términos para que dieran su traducción al español. Con el fin de que no dieran una que no tuviera nada que ver con el campo semántico elegido, se les dio la instrucción de que antes de comenzar la tarea se hicieran un panorama global de los términos. Para estar seguros de que se traducían en relación al campo semántico *Expert Performance*, se pidió a los sujetos que a parte de la traducción dieran una breve definición, en español, de cada una de las palabras. Además, dicha definición podía suplir la falta de vocabulario que a veces se tiene a la hora de traducir. Ya que si no hubiera sido por la definición, en varias ocasiones la persona encargada de corregir la prueba no se habría dado cuenta de si estos estudiantes realmente traducían y entendían correctamente el término dentro de un contexto determinado. Me refiero a términos como *savants* en el que muchos estudiantes dieron como traducción 'sabios', y para el cual la definición fue esencial a la hora de puntuarlo. Dicha prueba se evaluó de 0 a 1, se asignó 0,75 a la traducción del término y 0,25 a la definición. Las dos fases fueron corregidas por la misma persona, y siempre puntuando el mismo ítem en todos los cuadernillos antes de comenzar con el siguiente. Esto se hizo así para evitar el efecto halo, al que hace referencia Diekhoff (1983), en caso de que se puntuara estudiante por estudiante. Como ayuda se contó con un diccionario de Inglés-Español/Español-Inglés, con un monolingüe en inglés, y con un diccionario ideológico. Una vez evaluados todos los términos se halló la puntuación media por fase y por grupo para poder ser comparadas entre sí.

2.4.2. *Valoración de los términos*

Aparte de la traducción de los términos se le pidió a cada sujeto que estimara numéricamente el grado de relación entre cada pareja de conceptos. Se les dio una hoja de instrucciones, en la cual se les decía que el propósito era obtener su impresión general sobre la relación entre los ítems. Además, se daba una breve explicación sobre juicios de proximidad o de relación, y los detalles mecánicos sobre cómo introducir las puntuaciones en el ordenador Macintosh. Los 26 términos, los cuales aparecían en

la hoja de instrucciones para que los sujetos fueran estableciendo criterios de valoración, también se podían ver en la pantalla antes de dar comienzo a la tarea. Dicha tarea consistía en juzgar la relación existente entre todas las parejas posibles formadas con los 26 términos. Es decir, los sujetos tenían que valorar la relación o semejanza de 325 pares de conceptos (26 términos tomados de dos en dos) durante la sesión. En la pantalla del ordenador, debajo de cada pareja de términos, aparecía una línea horizontal con la palabra *low* a la izquierda y *high* a la derecha. Se hacía para señalar, dependiendo de si los conceptos estaban poco relacionados o mucho, hacia qué lado del teclado numérico debían presionar. Se pretendía evitar así confusiones. Indicaban su juicio (del 1 al 9), sobre la relación de cada pareja, presionando una tecla numérica. Si los conceptos no se relacionaban en absoluto debían presionar el '1', mientras que si pensaban que los conceptos estaban muy relacionados debían presionar el '8' o el '9'. Si se confundían podían volver a la pareja anterior, y una vez introducida la nueva puntuación la siguiente pareja aparecía automáticamente. El procedimiento continuaba hasta que todas las parejas se habían presentado de modo aleatorio, teniendo una duración media que oscilaba entre 25 y 30 minutos por tarea.

Una vez obtenidos los datos con la valoración original dada por los estudiantes se hizo una matriz simétrica triangular de 26x26 por sujeto. También se computó la matriz media del grupo experto, la cual sería más tarde utilizada para compararla con la de los estudiantes en las distintas fases y grupos. A parte de la medida de proximidad, la cual como ya se ha mencionado se obtiene a través de las valoraciones originales, se halló el valor de similitud para cada uno de los estudiantes y fases en relación al grupo experto; y posteriormente la media en ambas medidas para cada fase.

3. RESULTADOS

Para establecer la comparación, entre las fase previa y la fase posterior al estudio, en todo momento se tuvo como referente la estructura del grupo experto. Así, con las valoraciones del panel experto se halló la media con el fin de obtener una sola red ideal de referencia con la cual comparar las redes de los estudiantes. Para ello se hizo una correlación de las valoraciones originales de cada uno de los estudiantes con las del promedio experto. Además de la medida hallada con las valoraciones originales (proximidad) se halló también la medida de similitud. Para ello se siguieron pasos parecidos a los de antes, pero en este caso se sometieron las valoraciones originales al Pathfinder. Posteriormente se halló la media de las puntuaciones de proximidad y de similitud por grupos y fases. Se tomaron esta dos medidas, junto con las de la puntuación en la traducción de los términos, con el fin de poder comprobar si de una etapa a otra se habían producido cambios en la traducción y en la organización de los términos. Se espera que, si se produce el aprendizaje hipotetizado, dadas las diferentes investigaciones en el campo (Cooke y McDonald 1986; Goldsmith *et al.*

VALORACIÓN DEL APRENDIZAJE EN LA TRADUCCIÓN

1991; Gonzalvo *et al.* 1994; Johnson *et al.* 1995), el índice de similitud se muestre como la medida más adecuada para captarlo. Ya que en esas investigaciones, evaluando distintas medidas entre las que se hallan las de proximidad y similitud, conceden una ventaja estructural a las soluciones halladas con el Pathfinder.

Al comparar la estructura cognitiva de los estudiantes en las distintas fases y grupos, cuyo referente es el ideal experto, se intenta mostrar que se ha producido un aprendizaje por el estudio. Dado que el aprendizaje implica la reestructuración de la representación mental dentro de un campo de conocimiento, la representación estructural de la fase Pre I y Post I debería de cambiar y hacerse más similar a la de los expertos a medida que se produce el aprendizaje (Gonzalvo *et al.* 1994). Sin embargo, la diferencia que se esperaba en la organización estructural, representada por la medida de similitud, no se produjo. De hecho, las medias son exactamente iguales en las dos fases del grupo Pre/Post-estudio. En las otras dos medidas: traducción y proximidad, se apreció una mayor puntuación en las medias, como esperaba, para las fases Post I y Post II (ver tabla 1).

Grupo y fase	Traducción (#)	Proximidad (\$)	Similitud (&)
G. Pre-Post estudio (N=29)			
Fase Pre	0,604	0,267	0,133
Fase Post I	0,937	0,317	0,133
G. Post-estudio (N=25)			
Fase Post II	0,878	0,310	0,137

Media de las puntuaciones por grupo y fase. # Máximo = 1 ; \$ Rango = gran similitud (1) a poca similitud (0); & Rango = poca similitud (0) a gran similitud (1)

Tabla 1

Una vez establecido en nivel alfa en 0,05 (es decir, se está dispuesto a admitir un error del 5%) se realizaron pruebas *t* entre las medias en las distintas fases. Resultando significativas únicamente las que tenían que ver con las puntuaciones de la traducción y con la medida de proximidad. Así, la prueba *t* de Student, entre las dos fases experimentales del grupo Pre/Post-estudio, fue $t(28) = 6,03, p < 0,0001$ para la traducción, y $t(28) = 2,07, p < 0,04$ para la medida de proximidad.

En contra de lo esperado, la diferencia entre las medias en la fase Post I y Post II de la traducción fue significativa: $t(52) = 2,20, p < 0,03$. Dado que tanto el grupo Post I como el Post II habían recibido el tratamiento de la variable independiente, no se esperaba que entre ellos hubiera diferencia significativa en ninguna de las tres variables dependientes. La diferencia significativa en la medida de traducción indica un efecto, en esta tarea, producido por la repetición de la misma prueba.

Los resultados en los valores de similitud no captan los cambios estructurales que se presuponían desde la fase previa al estudio (Pre) a la fase posterior a dicho estudio (Post I). El que no haya cambios en la estructura conceptual de los estudiantes, de una fase a otra, indica que no han conseguido asemejarse al comportamiento del grupo experto. Por lo tanto, no se logra demostrar la tesis aquí propuesta de que la organización del conocimiento cambia debido al aprendizaje producido por el estudio. Pues no hay que olvidar que es precisamente a través de la medida de similitud, hallada con el Pathfinder, como se han captado cambios estructurales (Gonzalvo *et al.* 1994; Johnson *et al.* 1995). Los resultados obtenidos en las pruebas *t* y en las medias obtenidas (tabla 1), ascendentes en las distintas fases de la traducción y proximidad y prácticamente iguales en las distintas fases en los valores de similitud, más bien indican que la traducción y la organización de ese material no se producen de ningún modo de forma simultánea. Necesitándose poco tiempo para poder aprender los términos en una lengua extranjera, y mucho más para poder asimilarlos y organizarlos.

4. DISCUSIÓN

En relación a los datos hallados en esta investigación, dos son los puntos sobre los que hay que reflexionar y sobre los que no se ha dado explicación. Uno es el hecho de que la diferencia entre las medias en la fase Post I y Post II, en la traducción, sea significativa. Lo cual parece indicar que se ha producido un cambio debido a haber realizado la prueba con anterioridad. Si tenemos en cuenta la naturaleza de esta prueba, no es ilógico pensar en un cierto aprendizaje. El que en las otras dos medidas no se halla producido parece indicar una mayor sensibilidad a la repetición en la prueba de traducción. Es decir, el hecho de que la prueba fuera la misma en la fase Pre y Post I produjo un aprendizaje (conocido normalmente como *testing effect*). Sería preciso controlar este factor de forma más cuidadosa, en posteriores investigaciones, elaborando con el mismo tipo de material dos pruebas equivalentes en cuanto al nivel de dificultad.

El segundo punto sobre el que habría que meditar, es por qué hay un cambio significativo de una etapa a otra del aprendizaje en la traducción y en la medida de proximidad, y no en la medida de similitud. Se puede decir que se ha producido un aprendizaje, pero no el suficiente para que la estructura conceptual cambie. Es probable que las dos primeras medidas impliquen un aprendizaje más superficial, mientras que la última implicaría más bien un cambio estructural en la organización. Conclusión, que de replicarse en posteriores experimentos, sería de gran importancia y digna de ser considerada a la hora de evaluar el aprendizaje producido en un idioma, en este caso del inglés como lengua extranjera. Lo que nos llevaría a desechar nuestras intuiciones, muchas veces erróneas, y a dejar de evaluar a través de pruebas o

exámenes tradicionales que poco tienen que ver con un auténtico aprendizaje. Habría, pues, que volver la vista a trabajos experimentales, para así poder estar seguros en cada momento de qué es exactamente lo que se está enseñando y cómo se está evaluando.

BIBLIOGRAFÍA

A. Diccionarios

Lewis, N. 1978. *The New Roget's Thesaurus*. New York: Berkley.

Smith, C. 1971. *Collins Spanish-English English-Spanish Dictionary*. Glasgow: Collins.

Webster's New Collegiate Dictionary. 1980. Springfield: Merriam.

B. Referencias teóricas

Cooke, N. M., y J. E. McDonald. 1986. "A formal Methodology for Acquiring and Representing Expert Knowledge". *Proceedings of the IEEE* 74, 10: 1422-1430.

Diekhoff, G. 1983. "Testing Through Relationship Judgements". *Journal of Educational Psychology* 75, 2: 227-233.

Ericsson, K. A., y N. Charness. 1994. "Expert Performance. Its Structure and Acquisition". *American Psychologist* 49, 8: 725-747.

Goldsmith, T. E., P. J. Johnson, y W. H. Acton. 1991. "Assessing Structural Knowledge". *Journal of Educational Psychology* 83, 1: 88-96.

Gonzalvo, P., J. J. Cañas, y M^a. T. Bajo. 1994. "Structural Representations in Knowledge Acquisition". *Journal of Educational Psychology* 86, 4: 601-616.

Johnson, P. J., T. E. Goldsmith, y K. W. Teague. 1995. "Similarity, Structure, and Knowledge: A Representational Approach to Assessment". *Cognitively Diagnostic Assessment*. Eds. P. D. Nichols, S. F. Chipman y R. Brennan. New Jersey: Lawrence Erlbaum. 221-249.

Preece, P. F. W. 1976. "Mapping Cognitive Structure: A Comparison of Methods". *Journal of Educational Psychology* 68, 1: 1-8.

Schvaneveldt, R. W., F. T. Durso, T. E. Goldsmith, T. J. Breen, N. M. Cooke, R. G. Tucker y J. C. De Maio. 1985. "Measuring the Structure of Expertise". *International Journal of Man-Machine Studies* 23: 699-728.

APÉNDICE

Términos utilizados en el experimento:

Ability
Level
Acquisition
Master
Coaches
Processes
Development
Performance
Domain
Practice
Expertise
Skill
Knowledge
Goal
Gifts
Recall
Talent
Experience
Stimuli
Improve
Training
Speed
Attain
Prodigies
Cognitive
Savants