

## Para una nueva generación de diccionarios del alumno

PETER JAN SLAGTER  
Universidad de Utrecht

Este texto es la versión castellana del capítulo 7 de mi tesis Slagter, P.J. (2000), que redacté en inglés y que defendí en inglés y en holandés, en diciembre del año 2000, en la Universidad de Utrecht, Países Bajos.

Este texto explora alguna de las aportaciones que ofrece la tecnología educativa a la tarea de crear utensilios flexibles y amoldables a la etapa en la que se encuentre el alumno en cierto periodo de su aprendizaje. Concretamente nos proponemos sintetizar aquí los resultados de una serie de experimentos realizados a fin de crear un recurso autónomo para aprender el vocabulario de la lengua meta, en este caso el castellano, para estudiantes universitarios holandeses.

Estudios del campo de la psicología, y sobre todo los que enfocan al aprendizaje de vocabulario, afirman que las maneras en las que los sujetos procesan la información puede ser más importante que la mera intención o voluntad de aprender, mientras que estudios procedentes del campo de la enseñanza de primeras lenguas o de lenguas extranjeras insisten cada vez más en las posibilidades que parecen ofrecer las nuevas tecnologías. En el campo de las lenguas extranjeras gran parte del léxico puede "almacenarse" sin que haya intención de memorizar los vocablos en cuestión. Pero precisamente debido a considerables decepciones respecto a los beneficios del aprendizaje implícito hay quien se pregunta qué hacer con grandes cantidades de palabras que no se incorporan sin esfuerzos conscientes por parte del alumno.

Las limitaciones intrínsecas a los horarios de clases y la brevedad de muchos cursos nos obligan a explorar vías para fomentar que los alumnos aceleren el proceso de asimilación de vocablos, sobre todos los que son un tanto menos frecuentes, y el de construcción de redes semánticas propias de la lengua meta. La pregunta que pensamos desarrollar en esta contribución es básicamente ésta: los diccionarios electrónicos disponibles en línea pueden ayudar a los alumnos a resolver problemas de interpretación y traducción urgentes en determinados momentos, a la hora de leer o redactar textos, pero ¿hasta qué punto y de qué manera podríamos adecuarlos al nivel, se supone que en crecimiento constante, de los conocimientos léxicos de los alumnos?

Cuando los alumnos leen o escriben y consultan un diccionario, éste puede resolver dudas inmediatas, pero no es evidente que la actividad de buscar una traducción concreta, aplicarla a la tarea de ese momento, contribuya a almacenar a largo plazo ese vocablo nuevo, o reencontrado, al diccionario interno del alumno. Creemos que sólo se pueden hacer pocas

cosas, o quizás nada más que una sola cosa a la vez: o escribes, o lees, pero no puedes dedicar parte de tu atención consciente también a ampliar tu léxico al mismo tiempo de realizar las tareas que tienes entre manos. Incluso con la información lexicográfica al alcance de la mano, bien claramente impresa en la página correspondiente del diccionario, hemos visto que las prisas o la falta de destreza llevan a presentar traducciones, por no decir más, sorprendentes. Cualquier profesor de traducción reconocerá ejemplos como éstos: '*What a start you just gave me!*, traducido por *¡Qué comienzo me diste!*, y en vez de '*de vlag strijken*' ('*arriar la bandera*') nos encontramos con '*planchar la bandera*' que aparece destacadamente como colocación usual en el lema correspondiente a '*vlag*'. Las autoras de este despiste luego aclararon que habían buscado precisamente *vlag* -> *bandera*, pero que no se habían parado en que *strijken* podía ser parte de una colocación usual en holandés. *Strijken* es *planchar*, no les cabía duda. El problema, evidentemente, es más complejo que la simple pereza. También influye el hecho de que los hablantes nativos de una lengua tampoco tengamos conocimientos ilimitados de nuestra propia lengua.

Muchos profesores recomiendan el uso de diccionarios, y con razón. Un sector importante de los alumnos, sin embargo, defiende la posición opuesta: si a pesar de usar diccionarios cometemos cantidad de errores, ¿no será más sensato no hacerles caso? Lo que se tienen que plantear estos profesores viene a ser las siguientes preguntas: ¿de qué manera podemos conseguir que nuestros alumnos sigan dispuestos en realizar consultas al diccionario que tienen a mano? ¿Cómo podemos convencerlos de que los errores se deben o bien al hecho de que se han encontrado con información incompleta, deficiente, en resumen, ineficaz, o bien sus propias acciones de búsqueda y captura, deficientemente realizadas a falta de mejores utensilios, o por descuido suyo, y esto -claro está- sin dar más pasto a las sospechas que ya tienen? ¿Cómo mostrar que el esfuerzo que invierten ellos en consultas al diccionario sigue siendo utilísimo y de gran rendimiento? ¿Cómo mostrar que las actividades que proponemos para combinar consultas con esfuerzos para almacenar esas nuevas palabras –sirviéndonos de una tecnología y métodos nuevos, novedosos y extraños– son inversiones sensatas de alto rendimiento y que los resultados de sus esfuerzos se verán a corto y a largo plazo. ¿Cómo demostrarles que esas nuevas palabras han pasado a formar parte de su memoria permanente?

Nuestro estudio tiene tres objetivos: (1) demostrar que la configuración de las tareas afecta la conducta del alumno; (2) analizar posibles efectos variables de los tipos de información lexicográfica disponibles en un entorno digital; y (3) presentar resultados, mostrar no sólo que los sujetos del experimento que proponemos almacenarán las palabras trabajadas en el mismo experimento, sino que también el experimento en su conjunto da pie a que los sujetos transfieran su modo de trabajar con las palabras meta, a otras que no se tratan para nada en el curso del experimento, es decir, probar que es la misma conducta cognitiva de los alumnos, enriquecida con estrategias nuevas, la que permitirá que sean capaces de asimilar

elementos léxicos que encuentren en nuevos textos. Que haya *transfer* más allá de lo meramente memorizado y que ese *transfer* esté en la base de su aprendizaje continuado.

La relevancia del proyecto radica en que puede aportar una base teórica y práctica para ofrecer ayuda en línea para tareas de comprensión y producción de textos, al mismo tiempo que se estimule la retención de los vocablos desconocidos. El objetivo final, es, pues, analizar la retención como función de ayudas lexicográficas diferenciadas.

## CONTEXTO

Este estudio arranca en 1995, durante una clase en la que se daban multitud de ejemplos del tipo de '*planchar la bandera*'. Antes de esas fechas adoptábamos una actitud comprensiva (si no viene en ningún diccionario, ¿cómo íbamos a echarles nada en cara a nuestros alumnos?), pero a partir de la publicación de diccionarios bilingües adecuados (Slagter 1992, dicho sin modestia), ante la persistencia de este tipo de errores, no se nos ofrecía otra salida que reconocer que servirse atinadamente de un diccionario es más complicado de lo que parece. Por lo tanto, a finales del año académico de 1996-1997 diseñamos un curso en el que a los estudiantes se les presentaba en pantalla un fragmento "auténtico" con una palabra subrayada que había que traducir de acuerdo con el contexto. La clave contenía todas las traducciones medianamente posibles tal y como venían en Slagter (1992). Al final de cada subtarea del curso (fragmento + palabra subrayada + instrucción de dar una traducción adecuada al contexto del ese fragmento), los estudiantes tenían que traducir el fragmento completo al holandés. El profesor (un servidor, por más que en este contexto pueda resultar ambiguo el término) analizaba las respuestas y las subsiguientes traducciones. En las clases "normales" se comentaban luego las variantes aceptables de las versiones holandesas. Una de las cosas cada vez más evidentes era que la instrucción "traduce de tal manera que un lector holandés no se dé cuenta de que se trata de un texto traducido" frecuentemente llevaba a situaciones en las que la única traducción razonable por fuerza tiene que recurrir a términos no incluidos en los lemas del diccionario, no por deficiencias de éste -me apresuro a aclarar- sino porque los contextos imponen sus criterios y el grado de libertad que hay que tomarse. Los lexicógrafos ofrecen materiales crudos.

En esos años -y qué rápidamente pasas de una época a otra -no disponíamos de software que permitiese abrir más de una ventana a la vez. Trabajar en un texto de lectura presentado mediante ordenador y abrir otra pantalla con información lexicográfica era imposible, razón por la cual hemos usado un formato CALL (Computer Assisted Language Learning, o bien ELAO). Decidimos recurrir a una simulación CALL de un diccionario bilingüe e incorporar toda la información que éste normalmente hubiera ofrecido. Esta simulación tenía por objetivo proporcionar información sobre los consejos que pudiéramos dar a nuestros alumnos a partir del momento en que importantes y nuevos recursos informatizados

estuvieran a su alcance y que ojalá tuviesen en cuenta su nivel de desarrollo y no los aturdiesen con información lexicográfica que a todas luces tardaría bastante en hacerse relevante para sus tareas de lectura y redacción. Todo el experimento se realizó como si los materiales experimentales fuesen parte del programa normal de clase. No hubo necesidad de reclutar a voluntarios. El experimento lo hicimos con todos los alumnos del curso.

En este estudio, pretendíamos obtener dos tipos de datos:

- Sobre procesos de datos, las respuestas de estudiantes a tareas, y
- Mediciones sobre el producto, diferencias entre tests preliminares, intermedios y finales.

Los datos (a) son necesarios para saber en qué medida las tareas influyen en la ejecución de las mismas, y (b) pretende contestar a preguntas en torno al diseño de materiales. Sin esta combinación de datos no sabríamos qué les estábamos infligiendo a nuestros alumnos..

#### PROBLEMAS BÁSICOS EN ELAO, Y A LA HORA DE CREAR TAREAS

Nos referiremos a cuatro tipos base que todos contienen los siguientes campos con sus correspondientes acciones y efectos, pero difieren en cuanto a los contenidos de sus pantallas de ayuda. En el siguiente ejemplo, cada uno de los campos representa una pantalla nueva. El texto que se ve en cada caso viene bajo '*ejemplo*' y lo que experimentan los alumnos se detalla en la última columna. Éste es el formato de '*mal, intenta otra vez*', la variante NF, *No Feedback*. '#1' significa que sólo se reconoce como respuesta correcta una copia exacta de '*bestraft*', o '*gestraft*'. Si la clave incluyera '*straf*' (*castigo*, en holandés), el programa aceptaría sin más miramientos cualquiera de las siguientes secuencias, incorrectas pero de holandés comprensible: '*gestrafd*', '*strafen*', '*strafde*' y otros ejemplos de ortografía deficiente, aceptadas mientras apareciera la cadena de letra de '*s-t-r-a-f*'.

CAMPOS	PANTALLA	ACCIONES Y EFECTOS
Pregunta	.pregunta ¿Cómo se traduce la siguiente palabra? sancionado	El estudiante ve la pregunta y escribe la respuesta en la casilla. Luego le da al Enter
Clave	.respuesta1 bestraft, gestraft, #1	La respuesta se compara con la clave
Feedback	.feedback Tu respuesta no es correcta. Intenta otra vez.	El <i>feedback</i> negativo aparece a raíz de la respuesta incorrecta.
Respuesta correcta	.exit sancionado = bestraft Continúa con la pregunta siguiente.	La evidencia positiva aparece después de la respuesta inmediata o retrasada del alumno. Para las palabras acertadas esto es un refuerzo positivo, para las desconocidas, especialmente para las palabras que el alumno no ha acertado incluso tras diversas tentativas, esto es <i>feedback</i> negativo.

Tabla 1. Diseño básico de interacciones en ELAO (pregunta, clave, *feedback*, y respuesta correcta)

El campo '*feedback*' cambia con el contenido y lugar en el desarrollo del estudiante en la interacción entre el estudiante y la pregunta. En este ejemplo, el *feedback* negativo se da en la forma de un mensaje de error sin definir. Creemos que es razonable suponer que este tipo de *feedback* tan uniforme, general y poco adaptado a la pregunta en cuestión, no ayudará mucho al alumno. Sin embargo (Schachter 1991), adopta una posición mucho más positiva. Sobre este tipo de mensaje dice:

*'The reader who has never been a subject in a problem solving experiment might focus on the times you have had to learn how to do something on the computer with no one around you to instruct you. You receive both positive and negative feedback in those situations, and, for the most part, you are able to make use of that negative feedback to learn both what is and is not allowed as you try to achieve your goal. The information 'no, it's not that way' is tremendously valuable in such a task, because it cuts one's hypothesis space, successively narrowing the set of possible hypothesis to be tested'.* (Schachter 1991: 93).

El funcionamiento de una interacción de este tipo lo revelan los '*logfile*s' de los alumnos, que contienen cada letra o signo que haya picado. Estos ficheros recogen todas las respuestas dadas en sucesivas vueltas y los presentan en columnas que contienen las respuestas dadas

en todo el tiempo que el alumno dedique a la pregunta en cuestión, al mismo tiempo que da la puntuación tal y como se desarrolla al paso del alumno por la interacción, desde la primera hasta la última tentativa. Así, estas tablas nos permiten analizar el efecto de los diferentes tipos de *feedback*. En el ejemplo simplificado que presentamos a continuación, la puntuación es de 00 a 10. Un 10 en la primera columna significa que el alumno contestó correctamente, sirviéndose de conocimientos existentes que le permitieron dar una de las posibles traducciones previstas y correctas. Un 10 en la segunda columna demuestra que probablemente el alumno hiciera lo que describió Schachter (1991: 93): se dio cuenta de que el conjunto de hipótesis había quedado reducido y volvería a analizar la pregunta, y puede que buscara de verdad en su léxico interno, después de lo cual puede que corrija su anterior optimismo, falta de atención, supere su ingenuidad y otros distractores, y que dé la respuesta correcta. Ésta aparece en segundo lugar en la serie de respuestas dadas. Un 10 en la columna número 3 significa que el alumno dio la respuesta correcta después de ver la clave, que es la evidencia positiva contenida en el campo .exit, y después de analizar el *feedback* negativo dado al cabo de la primera y segunda tentativa. Esto es lo que pasa después de completar todas las preguntas de la serie y al volver a las preguntas incorrectamente hasta ese momento.

ESTUDIANTE	1	2	3	4	
1		10	10	10	bestraft
2		00	10	10	gesanctioneerd   bestraft
3		00	00	10	'no sé' *   'sigo sin entender' *   bestraft

\* En la columna 4 se recoge todo lo que escriban los alumnos en las sucesivas vueltas; las barras indican al paso de una vuelta a otra; los alumnos no pueden pasar adelante a la pantalla siguiente sin por lo menos picar algo.

Este formato, entre paréntesis, resuelve uno de los más arduos problemas en experimentos de aprendizaje de vocabulario, el de determinar el conocimiento disponible antes del experimento y evitar que se contaminen los datos por palabras disponibles en la interlengua del alumno sin que los investigadores hayan podido localizar ese conocimiento. En este formato, el tratamiento empieza después de la primera vuelta, y para medir el efecto del tratamiento, podemos comparar los conocimientos recogidos en la primera columna con los efectos del tratamiento que se hacen visibles en la segunda columna (las respuestas dadas después de aparecer el mensaje de '*mal, intenta otra vez*', mientras que la retención inmediata –sin salir de la sesión– se evidencia en la tercera columna ('*al principio no sabías, no conseguiste aprovechar la segunda oportunidad, pero ahora parece que aprendiste el significado que me acabas de dar, y lo diste a base del feedback que te he dado*').

Los siguientes párrafos describen sucintamente las variedades de evidencia positiva del

formato CALL/ELAO por el que hemos optado.

### TRATAMIENTOS, PROCESOS, PANTALLA DE AYUDA Y OPERACIONES COGNITIVAS

En cuanto a las diferencias entre las pantallas de ayuda, es importante destacar que el campo .exit, el último de la sucesión de pantallas, de todos los tipos de *feedback*, contiene toda la información esencial. Las interacciones del tipo '*mal, intenta otra vez*' no implican que al estudiante le dejemos perplejo sin tener idea de lo que no ha sabido contestar. De la misma manera, los otros tipos siempre acaban ofreciendo toda la información relevante. En el fondo, las diferencias que podamos identificar al pasar los estudiantes por las tareas son mínimas. En las del tipo '*mal, intenta otra vez*', tardará un poquito más en llegar a la respuesta correcta; en las del tipo 2, acto seguido de dar una respuesta errónea. al estudiante se le invita a interpretar una definición e identificar un equivalente en lengua materna; en las del tipo 3 los estudiantes tienen que encontrar la o las traducciones que vayan con palabras de varios significados; mientras que en las del tipo 4, las de las citas o fragmentos del texto original de donde viene la palabra en cuestión, la pantalla de ayuda invita a descifrar ese fragmento de contexto. Cualquiera de las interacciones y cualquiera de las pantallas de ayuda terminan dando la solución correcta. Lo que nos toca hacer es demostrar que la actuación revela varianza significativa a lo largo de curso.

Los estudiantes no pueden simplemente darle al botón Enter para pasar de una pantalla a otra para ver las respuestas correctas, a pesar de lo cual éstas se encuentran a dos pasos nada más. Podríamos pensar que en las preguntas con *feedback* del tipo '*mal, intenta otra vez*', el tiempo que se necesita para pasar de una pantalla a otra aumenta las expectativas y la receptividad de los alumnos ante esas respuestas correctas. Por otro lado, demasiados mensajes de este tipo pueden muy bien entumecer la mente de los sujetos y llevarlos a no prestar ninguna atención a las claves. Lo mismo podemos decir a propósito de las demás opciones; cuanto mayor la diferencia entre los conocimientos de acceso inmediatos o las destrezas de identificación y la dificultad de la pregunta en cuestión, tanto mayor será la curiosidad del alumno por saber la respuesta. La distancia entre lo sabido y lo que está por saber, lógicamente, puede también llegar a frustrar al alumno.

Cada uno de estos formatos de *feedback* tiene sus propias peculiaridades cognitivas. El alumno que sepa utilizar una frase tomada de la fraseología lexicográfica, o una serie de frases relacionadas, es capaz de abstraer la estructura de una locución concreta, como lo hace a base del contexto de la tarea de comprensión lectora o de producción escrita que le toca trabajar, mientras que sabrá hacer eso mismo en el caso de las frases ejemplo de la fraseología de un diccionario. Esto quiere decir que analizará material polisémico y sus estructuras a partir de los ejemplos que encontrará en la fraseología y lo confrontará con la tarea que tiene entre manos. Estos pasos parecen nimios pero requieren dotes de

interpretación no por casi implícitos menos considerables antes de poder establecer el equivalente idóneo para el contexto de la primera o segunda lengua.

Las frases reunidas en la fraseología están ahí para ofrecer información relevante. Un verbo, por ejemplo, se presenta en una frase de tal manera que se vean los sujetos y complementos gramaticales usuales, de tal manera que no solo se refuercen conocimientos de la primera lengua en torno a los elementos de la oración, sino que también se marquen contrastes cuando así se requiera. Todo esto pasa sin que alumno ni profesor se den cuenta cabal y explícita de los procesos en juego. No está de más destacar aquí que estamos hablando de complejas actividades cognitivas. Por tradición suponemos que leer textos auténticos e inferir el significado de palabras desconocidas o de acepciones nuevas de palabras ya más o menos conocidas –y más si se trata de palabras indispensables para comprender el texto en cuestión- sea una actividad beneficiosa para almacenar esos vocablos y esas nuevas acepciones en el léxico permanente del alumno. Al mismo tiempo, es decepcionante tener que reconocer hasta qué punto leer un texto es una actividad y memorizar nuevas palabras otra distinta. La atención humana no se duplica porque así nos parezca normal y deseable a los profesores. Nuestra hipótesis es que las citas de textos auténticos no amañados serán menos explícitas y menos propicias que las frases tomadas de un diccionario con buenas fraseologías, simplemente porque éstos están redactados con otros criterios que los que rigen para textos escritos para aquellos lectores que se enfrentan a textos en su primera lengua. En estos últimos lo anormal sería incluir explícitamente la información que necesita nuestro alumnado.

El usuario de un diccionario bilingüe se supone que active conocimientos de su primera lengua a fin de completar información implícita (una vez identificada una palabra como verbo, sustantivo o adjetivo automáticamente activa esquemas de la primera lengua que no se incluyen en diccionarios bilingües), pero éstos son conocimientos no fácilmente accesibles para los usuarios de diccionarios monolingües. Por consiguiente, creemos que las definiciones redactadas en lengua meta presentan problemas adicionales al usuario extranjero, ya que describen las palabras en vez de ofrecer una o más traducciones directamente aplicables. Diríamos que parecen responder a preguntas no formuladas.

Los perfiles semánticos con o sin paráfrasis de los elementos recogidos en el perfil semántico, frases ejemplo monolingües y definiciones redactadas en lengua meta son los aliados naturales del usuario que opera en su propia lengua, pero pueden no ser más que falsos amigos para nuestros alumnos.

Como primer paso, tendremos que demostrar que los estudiantes reaccionan de forma diferente ante diferentes tipos de información. También habrá que demostrar que los estudiantes querrán aprender con el sistema que utilizamos, que conseguiremos que se



propongan aprender con estos materiales, que podremos hacerles jugar con ellos, que respondan al trapo y que intenten vencer el sistema, como si fuera un videojuego.

## DEFINICIONES

El modelo que describíamos arriba, el formato de '*mal, intenta otra vez*' sólo sirve si se interpreta como estímulo para perseverar: '*chico, te toca hacer algo, da media vuelta y hurga tus conocimientos: puede que te haga falta analizar palabras, comprobar cognados, o guardarte de falsos amigos*'. Para que este tipo de actividad tenga éxito, los conocimientos relevantes tienen que estar disponibles en la memoria del alumno. Si, como en este caso, el alumno no conoce la palabra, no hay forma de que analice *sancionado*. Saber que *sanción* es algo así como '*castigo*' y que aplicar la "regla" de que *sanción* + el marcador de participio -ado puede llevar a '*castigado*'. Si no conoce *sanción*, hará falta darle más evidencia positiva. ¿Qué información puede servir de ayuda, evidencia o *feedback*? El siguiente formato se sirve de una o varias definiciones en lengua meta (Tabla 2). Por supuesto que es discutible la decisión de restringir la información dada. Puede que sea mejor darle toda la información que el estudiante encontrará en la versión impresa o la versión electrónica completa de la palabra buscada. Cuando diseñamos la versión original de esta investigación (1997) no hemos optado por esta solución, como tampoco incluimos todas las frases ejemplo de la fraseología teóricamente disponibles (Tabla 3).

CAMPO	PREGUNTA	ACCIONES Y EFECTOS
Pregunta	.pregunta ¿Cómo se traduce la siguiente palabra? sancionado	El estudiante ve la pregunta en la pantalla, pica la respuesta en la casilla correspondiente y le da al Enter.
Clave	.respuesta1 bestraft, gestraft, #1	La respuesta se compara con la clave.
Feedback	.feedback Mira a ver si te sirve esta definición: castigar o poner una pena	El <i>feedback</i> negativo es implícito en la pantalla que ofrece la primera de tres definiciones disponibles en Vox-Alcalá.
Respuesta correcta	.exit sancionado = bestraft. Continúa con la pregunta siguiente.	La evidencia positiva aparece después de una respuesta inmediatamente correcta o correcta al cabo de varias vueltas: para las palabras inmediatamente reconocidas esto es un refuerzo positivo, para aquellas palabras desconocidas, y en especial las que no se hayan contestado correctamente ni inferidas en dos o más vueltas, esto es <i>feedback</i> negativo.

Tabla 2. Diseño básico de interacciones CALL/ELAO para definiciones.

La Tabla 2 contiene la información elemental sobre el *feedback* basado en definiciones. En este ejemplo, no hemos incluido las definiciones número 2 y 3, ya que no valen para el texto original y no se conforman a la clave requerida para incluir "citas" como tipo de *feedback*. Estas otras definiciones (2) *autorizar o aprobar un acto, uso o costumbre*; (3) *dar fuerza o carácter de ley*. Incluir estas opciones obligaría a aceptar '*gesanctioneerd*', '*bekrachtigd*', '*goedgekeurd*', '*erkend*' (además de las ya aceptadas '*gestraft*', '*bestraft*'), '*haber aplicado una sanción*', '*haber impuesto un castigo a*' en el campo .exit.

Las definiciones en lengua meta se suelen incluir sólo en diccionarios monolingües porque obviamente faltan otros medios para desambiguar los significados. Nosotros hemos incluido esta opción como una posibilidad no fantástica de ofrecer información porque no descartamos que en un futuro no demasiado lejano se puedan incluir en versiones electrónicas amoldables al nivel de los alumnos. Almacenar información lexicográfica electrónicamente en diccionarios en CD-ROM es un problema insignificante comparado con

el coste en papel de sus homólogos tradicionales. A medida que suba el nivel de nuestros alumnos puede que se beneficien de este tipo de información que es el que se observa en la vida normal es interacciones entre hablantes nativos con hablantes no nativos.

FRASES		
CAMPOS	PREGUNTA	ACCIONES Y EFECTO
Pregunta	.pregunta ¿Cómo se traduce la siguiente palabra? sancionado	El estudiante va la pregunta y pica la respuesta en la casilla correspondiente de la misma pantalla. Termina con Enter.
Clave	.respuesta1 bestraft, gestraft, #1	El programa compara la respuesta con las respuestas aceptadas de la clave.
Feedback	.feedback Mira a ver si te sirve este ejemplo: 'las autoridades municipales sancionaron al dueño del establecimiento con seis meses de suspensión de la licencia'	La información, tipo <i>feedback</i> negativo, es implícita cuando la pantalla de ayuda aparece con la primera frase de la fraseología de Vox-Alcalá. Por motivos idénticos a los explicados al pie de la Tabla 2, los ejemplos 2 y 3 del perfil semántico se suprimen en este diseño.
Respuesta correcta	.exit sancionado = bestraft. Continúa con la pregunta siguiente.	La evidencia positiva aparece después de una respuesta correcta inmediata o retrasada; para las palabras conocidas esto es refuerzo positivo, para palabras no conocidas, en particular en el caso de no acertar un par de veces, esto es <i>feedback</i> negativo.

Tabla 3. Diseño básico de interacciones CALL/ELAO para frases.

Hemos optado, sin embargo, por jugar con el grupo total de frases (Tabla 3) contenidas en los lemas del diccionario monolingüe por dos motivos: (a) en diccionarios monolingües y bilingües, acudiendo a las fraseologías, los alumnos se enfrentan a una serie de frases ejemplo, y frecuentemente su totalidad tampoco puede traducirse mediante un solo equivalente en lengua materna. Nada más natural en lo referente a palabras polisémicas; (b) en futuros pero posibles desarrollos tecnológicos puede que sea recomendable distinguir entre información masiva y otra más selecta, por ejemplo utilizando un paradigma heurístico que empiece pidiendo a los alumnos que la parte del perfil semántico que mejor vaya con el contexto de la lengua meta y permitiéndole luego que acceda a la fraseología que corresponda al significado elegido, por lo cual la serie de frases ejemplo se limitará a una

fracción de las contenidas en el lema.

Estas frases ejemplo hacen el papel de evidencia positiva (Tabla 3). Naturalmente, el alumno no acude a ellas sino después de dar una respuesta errónea, y cuando aparecen estas frases ese mismo alumno también se da cuenta de que representan *feedback* negativo ante su primera respuesta. La diferencia más grande entre los diccionarios impresos monolingües y bilingües es que éstos incluyen información en L1. Para nuestros propósitos es interesante que la información monolingüe permita comparar la calidad del *feedback* positivo. Decidimos no incluir la fraseología de los Van Dale, ni tampoco las traducciones holandesas, por obvios motivos.

QUOTES		
CAMPOS	PREGUNTA	ACCIONES Y EFECTOS
Pregunta	.pregunta ¿Cómo se traduce la siguiente palabra? sancionado	El estudiante ve la pregunta que aparece en la pantalla y escribe la respuesta en la casilla correspondiente y le da al Enter.
Clave	.respuesta1 bestraft, gestraft, #1	Se confronta la respuesta con las soluciones que tiene la clave.
Feedback	.feedback Mira a ver si te sirve la siguiente cita: 'El flamante nuevo director de TVE, Fernando López-Amor, inspector fiscal de profesión, fue sancionado por abuso de autoridad en su etapa de jefe de la Unidad de Inspección de la Administración de Hacienda de Madrid'.	El <i>feedback</i> negativo es implícito. Después de la respuesta incorrecta, en la pantalla de ayuda aparece una cita procedente del texto original.
Respuesta correcta	.exit sancionado = bestraft. Continúa con la pregunta siguiente.	Aparece la evidencia positiva después de una respuesta correcta inmediata o retrasada. Para las palabras conocidas o acertadas este detalle significa evidencia positiva, para aquellas palabras que el alumno no consiga identificar una o más veces o cuya identificación no consigue hasta que pase algún tiempo esto equivale a <i>feedback</i> negativo; ante palabras conocidas este campo significa refuerzo positivo.

Tabla 4. Diseño básico de interacciones en CALL para citas.

En clases de lenguas solemos trabajar con citas y con textos auténticos (ver Tabla 4), como lo son –en este caso– textos periodísticos. Se supone que éstas representen tareas de comprensión lectora normales, pero hay que destacar dos aspectos importantes: a) se suelen deducir los significados de palabras desconocidas a base de fragmentos muy cortos, sin recurrir a la cita completa ni mucho menos al texto íntegro cuando éstos se ofrecen; b) a nuestros estudiantes solemos recomendarles que lean muchos textos, que “hagan kilómetros”, para enriquecer su léxico (ver Tabla 4). Pero cabe preguntarse si emplear tanto tiempo leyendo textos auténticos puede ser todo lo eficaz que se pretenda. Al fin y al cabo, estos textos no se redactaron con alumnos no nativos en mente ni mucho menos para que éstos aprendieran nuevas palabras de la lengua meta. Tampoco ofrecen ayuda para desambiguar significados que el lector nativo o avanzado deduce del contexto.

#### INSTRUMENTOS PARA ELAO

Para que nuestro diseño pudiese funcionar tal y como queríamos, las palabras seleccionadas tenían que ser de baja frecuencia. Si hubiéramos utilizado palabras frecuentes, a los alumnos les hubieran resultado fáciles y no hubiéramos tenido la oportunidad de medir la eficacia de los distintos tipos de ayuda lexicográfica que pensábamos incluir. Además, resultó necesario reducir las preguntas a una sola palabra o un solo grupo de palabras. Más contexto nos hubiera podido obligar a revisar los textos originales y hubiera dado lugar a claves muy complejas. De esta misma manera, a la hora de acceder a la segunda oportunidad para contestar a la pregunta original, las tres variantes de ayuda lexicográfica ofrecidas podían demostrar sus diferentes grados de eficacia como instrumento de inferencia, mientras que la reacción del tipo '*mal, intenta otra vez*', cuando menos tenía la virtud de redirigir la atención del alumno a partes de la pregunta, raíz de la palabra, terminación, letras iniciales o combinaciones de letras presentes en alguna parte de la palabra meta.

En CALL, las claves incluían la información relevante en holandés, tomada del diccionario Van Dale Español-Holandés. “Relevante” apunta a una o más opciones que pudieran aplicarse a los contextos originales de los que reunimos en las series de interacciones CALL, preguntas, para cada sesión en el laboratorio. En las pantallas de ayuda, sin embargo, nunca incluimos contenidos del Van Dale Español-Holandés. Van Dale no incluye definiciones en lengua meta. Si hubiéramos incluido la fraseología de Van Dale con las correspondientes traducciones al holandés, al alumno no le hubiéramos podido exigir ningún tipo de esfuerzo mental que valiera la pena por deducir el significado en cuestión. Por este motivo, nos servimos de las definiciones y la fraseología del diccionario Vox-Alcalá monolingüe para estudiantes extranjeros (Alvar Ezquerro, 1995). En cuanto a las preguntas basadas en citas, no hemos tenido que hacer más que copiarlas de los textos originales. Con esta reducida serie de variantes creíamos poder comprobar si los significados se deducían, se almacenaban y estarían disponibles para futuros usos.

Pensábamos que las definiciones en lengua meta serían útiles por dos motivos. A la hora de preguntar a otra persona el significado de tal o cual vocablo, no es raro recibir una definición por respuesta. El diccionario monolingüe para alumnos extranjeros (Alvar Ezquerra, 1995) constituía una gran ayuda porque –caso de revelarse como ayuda positiva en este experimento que al mismo tiempo era un test de uso –implícito- de la ayuda que ofrece este diccionario de alumnos– podríamos aconsejar su uso a nuestros alumnos. Como ya dijimos, en aquellos casos en los que Vox-Alcalá ofrecía más de una definición o frases ejemplo, sólo hemos incluido aquellas que fueran bien con la versión de citas y el significado que habíamos encontrado en los textos originales.

La última categoría consiste de citas de textos auténticos, es decir, citas de textos de lectura que en todo caso muy bien podrían usarse en esa misma clase. De entre los más de 500 artículos, cada uno de ellos buen material para tratar en clase o para tareas de lectura independiente, hemos elegido los que contuvieran series de palabras poco frecuentes. Nos hemos servido de un programa de concordancia para preparar una lista de frecuencias, para luego quedarnos con los textos de mayor número de palabras dentro del mismo rango de frecuencia. Lo mismo hicimos para elegir las preguntas de las tres versiones del test. Nos hemos permitido trabajar con palabras relativamente poco frecuentes porque pensábamos que sería en esta área donde nuestra actividad didáctica podría ser beneficiosa para estos estudiantes intermedio/avanzados.

Si se tratara de textos normales de lectura, del tipo que se comprueba mediante preguntas de comprensión y no por preguntas de tipo sintáctico o morfológico, los alumnos probablemente ni se hubieran parado a pensar sobre su significado. Leer para comprender no hubiera sido problema. Pero al mismo tiempo, tampoco se pararían para desplegar actividades especiales destinadas a ampliar su léxico para futuras ocasiones o para uso productivo. Leer fiándose de estrategias deductivas es una digna actividad. Pero al mismo tiempo de dar sabios consejos del tipo '*lee, hijo, porque así aprendes más palabras*', los profesores creemos que ese rendimiento se produce sin esfuerzo consciente. Nosotros creemos -ya lo hemos dicho repetidas veces- que nadie puede hacer dos cosas a la vez, y si lo intenta, una de ellas sale perjudicada. En la fase de la carrera en que se encuentran estos estudiantes, no podemos permitirnos dejar al azar la construcción de un léxico más amplio y más sofisticado. De ahí la idea de intentar demostrar que -en ciertas condiciones- una tarea no excluye tan tajantemente la otra.

Los recuentos de frecuencia se basan en un simple (pero añadiríamos: sano) análisis docente. En el campo de los estudios hispánicos no abundan los recuentos de frecuencias. Además, tienen criterios de organización diferentes a los que se requieren hoy en día y los de actualidad más permanente, como por ejemplo Juilland y Chang-Rodríguez (1964), empiezan

a llamar la atención por el sesgo que los caracteriza. Creo que la mayoría de los profesores de lengua hoy en día selecciona textos por el interés que ofrecen y no por ajustarse a criterios de frecuencia. Tengo que destacar -para no dar lugar a voces de alarma- que mientras preparaba las interacciones CALL/ELAO (más de 12 capítulos, 32 nuevas preguntas en cada uno de ellos) no eran más de 4 o 5 los vocablos no incluidos en Van Dale ni en Vox-Alcalá o incluidos sin la información necesaria.

Limitando más aún la selección de textos, me quedé con los textos con mayor incidencia de palabras dentro de los márgenes de frecuencia adoptados, al mismo tiempo que volví a apurar la selección de temas de esos textos. Los así seleccionados, a mi juicio, tenían temas apropiados para estudiar en clase y palabras de utilidad y frecuencia razonables. Estos dos criterios eran necesarios porque no quería excluir un uso posterior en clase de los textos completos en una fase posterior a la de CALL/ELAO: manteniendo abierta esa opción a los estudiantes les podría dar la satisfacción de reconocer y volver a usar en ejercicios de *cued recall* (reactivación con estímulo o clave) por lo menos una parte de los vocablos que habían estado trabajando poco tiempo antes.

#### DICCIONARIOS MONOLINGÜES Y BILINGÜES COMPARADOS

Los diccionarios monolingües y los bilingües ofrecen diferentes tipos de información, se dirigen a distintos grupos de usuarios y recurren a diferentes métodos de procesar la información ofrecida:

DICCIONARIOS MONOLINGÜES	DICCIONARIOS BILINGÜES
<p>Los significados se explican mediante:</p> <p>Perfiles semánticos, con o sin resúmenes, paráfrasis.</p> <p>Definiciones en la lengua meta.</p> <p>Fraseología, clasificados por colocaciones o estructuras de oraciones.</p> <p>Recurren a los conocimientos de la primera lengua del usuario.</p>	<p>Los significados se explican mediante:</p> <p>Perfiles semánticos, más resúmenes y traducción de los elementos más importantes.</p> <p>(no tiene equivalente)</p> <p>Fraseología, clasificados por colocaciones o estructuras de oraciones, más traducción.</p> <p>Se aclara el significado recurriendo a ejemplos de uso, apoyado en los conocimientos de la primera lengua del usuario.</p>

Tabla 5. Funciones más importantes de diccionarios monolingües contrastados con diccionarios bilingües.

Esta distinción es arbitraria, sin embargo, a partir del momento en el que trabajar con varios tipos de información simultáneamente debe considerarse como buena práctica pedagógica. En entornos multimediales hay espacio virtual suficiente como para usar diccionarios monolingües y bilingües, al igual que diccionarios de sinónimos y tesoros de la lengua. Sin embargo, antes de incluir todos esos recursos es preciso saber mucho más sobre la calidad de la ayuda que de buena fe creemos ofrecerles a los alumnos de diferentes niveles de destreza.

En el presente diseño incluimos diferentes tipos de información tomados de diccionarios monolingües y bilingües, y lo hacemos así porque –como sugerimos antes– creemos que en entornos tecnológicos modernos resultará fácil combinar los dos tipos de información tradicionalmente separados por barreras pedagógicas y nacionales. Además, estos dos tipos de información permiten construir un conjunto coherente de recursos que hará más justicia a recientes criterios formulados con respecto al aprendizaje y la enseñanza de vocabulario. Llegado a esta fase de nuestra descripción puede que aún suene lejano, pero no podemos dejar de destacar que creemos que esta combinación de criterios y materiales puede llevar a nuevas conceptualizaciones sobre lo que se entiende por “diccionarios del alumno”.

#### INSTRUMENTOS PARA TESTS INICIALES, INTERMEDIOS Y FINALES

Además de los materiales CALL/ELAO, hemos preparado un test que pudiera servir como test inicial, intermedio y final. Este test comprendía 55 preguntas que se tocarían en el programa experimental y 57 preguntas con vocablos que no aparecerían en las interacciones. Incluir estas palabras nos permitiría analizar los eventuales efectos de *transfer* de estrategias. Los alumnos tenían que contestar a las 102 preguntas en un máximo de 60 minutos en las tres ocasiones de administrarse el test.

Las preguntas provenían del conjunto total de textos bajados de *El País* siguiendo el mismo procedimiento que hemos explicado arriba cuando nos referíamos a la selección de vocablos y citas para incluirlas en CALL/ELAO. El test consistía en fragmentos cortos con una palabra o un grupo de palabras subrayados que había que traducir al holandés reflejando el contexto y la variante más indicada. La clave de corrección respetaba respuestas posibles teniendo en cuenta los breves contextos ofrecidos, y lógicamente a los estudiantes se les había indicado que cuando se tratase de palabras de múltiples significados, tuviesen en cuenta el contexto, y no responder por cualquier variante lexicográficamente posible del vocablo en cuestión. No hemos pedido traducciones completas de los fragmentos, cosa que llevaría demasiado tiempo y no concordaba con nuestra intención de medir la capacidad de inferencia, más que la de dar traducciones adecuadas.

La serie de las 55 preguntas incluidas en las interacciones CALL esperábamos que resultara más y más conocida a lo largo del periodo de 4 meses que duraron los experimentos,



mientras que la otra serie, las 57 preguntas de control, sólo se verían en tres o cuatro ocasiones a lo largo del experimento. El intervalo entre una administración y la siguiente, sin embargo, considerábamos que era tan largo que apenas deberíamos temer por efectos de retención a causa de familiaridad con los tests. Contestar a 112 preguntas, si bien es verdad que al final del experimento la mitad procedía de citas que habían estado trabajando en el laboratorio, es como contestar a 112 preguntas individuales, tarea para la cual sólo podían basarse en el contexto inmediato de la cita. De hecho, sólo 55 preguntas podían reconocerse a base de los ejercicios CALL/ELAO, es decir, sólo las que se hubieran provisto de ayuda en clave de cita de contexto, y eso, además, sólo al final del experimento, cuando ya formaban parte de una serie de vocablos estudiados mucho más numerosa.

La cantidad de preguntas y la aparente falta de conexión entre el conjunto total de preguntas hacen que creamos que necesitamos temer un efecto de memoria de test de alguna importancia. Nos ha parecido éste un riesgo aceptable y el método adoptado nos ha parecido también preferible a construir tres juegos de preguntas diferentes sin tener ocasión de determinar sus características psicométricas. Las tres variantes del test –la variante 2 tenía las preguntas ordenadas de final a comienzo, numeradas de 107 para abajo– fueron administradas con seis u ocho semanas de intervalo. Para terminar, tenemos que dejar constancia del hecho de que tampoco les dimos la clave ni *feedback* de contenido a sus resultados a los estudiantes en cuestión que todos tuvieron la paciencia y confianza suficiente como para fiarse de nosotros.

Nuestras expectativas en torno al segundo juego de datos (puntuación en tests iniciales, intermedios y finales) son las siguientes: observaremos una diferencia significativa en retención, medida en términos de crecimiento, puntuaciones más altas que van del test inicial, por el intermedio hasta el final; y en la referente a tipos de tratamiento y efectos del *feedback* en evidencia positiva en las pantallas de *feedback*, el orden de aportación creciente pronosticamos que irá de '*mal, intenta otra vez*', a definiciones, por la fraseología hasta citas de los textos originales. Hemos comprobado una diferencia significativa en cuanto a la capacidad de inferir significados entre los tres tests (inicial, intermedio y final) entre las palabras incluidas en las interacciones y aquéllas que no fueron incluidas en el diseño experimental. Lógicamente, las palabras seleccionadas para formar parte del test fueron seleccionados de entre los mismos textos, la misma zona de frecuencia dentro del corpus de ítems seleccionados, y hemos hecho todo lo posible por evitar que destacáramos cualquiera de estas palabras durante la totalidad de curso. No obstante, puede que algún estudiante haya reparado fuera del contexto de esta investigación alguna o varias de estas palabras. Al mismo tiempo, consideramos que es muy poco probable que ese estudiante haya trabajado conscientemente esas palabras. Reconocer preguntas del test incluidas en el formato CALL/ELAO parece posible (pero en ninguna sesión CALL los participantes llegaban a la cota del 100% de aciertos), pero recordar palabras no incluidas en CALL/ELAO, por ejemplo

acordarse de haberlas visto en CALL durante lecturas de otro tipo era, repetimos, muy poco probable, ya que los temas elegidos no procedían del campo de la filología hispánica. La única oportunidad para establecer paralelos y reactivar conocimientos la darían las clases de lingüística y literatura, donde precisamente estos tipos de textos de información general no estaban en el menú.

Nosotros creemos que los aumentos de puntuación en lo referente a palabras no experimentales pueden explicarse como señal de cambios en la destreza de "adivinar a base del contexto". Así, las puntuaciones más altas indicarían que los estudiantes daban cada vez mejor no sólo en los significados de palabras estudiadas sino que también en las que jamás habían visto, siempre en el espacio de los 60 minutos que tenían para responder a las diferentes versiones de los tests. Las altas correlaciones entre los aciertos procedentes de las varias maneras de tratar las palabras del experimento y las palabras no tratadas ni provistas de ninguno de los tipos de ayuda indican, a nuestro juicio, un incremento en capacidad deductiva, efecto del experimento en su conjunto.

#### PRIMER ESTUDIO (1997)

Como primer paso, necesitamos demostrar que las tareas CALL/ELAO afectan la manera de procesar las palabras ofrecidas. En el diseño de 1997 hemos utilizado una especie de rotación en los *feedback* al interior de cada capítulo. La pregunta número 1 era del tipo '*mal, intenta otra vez*', la pregunta 2 ofrecía una definición, la 3 brindaba fraseología y la 4 recogía una cita del contexto original.

La Tabla 6 recoge la distribución de los tipos de *feedback* en las pantallas de ayuda. Intentaremos demostrar que las reacciones de los estudiantes varían con el tipo de *feedback*. Este estudio, como ya hemos dicho, se efectuó en una clase normal, pese a que dispusiéramos de ordenadores individuales, detalle que impidió que efectuáramos una investigación experimental. En su lugar, el hecho de servirnos de una clase normal y corriente, y utilizar los materiales experimentales como parte integral de las clase, obligó a montar un diseño cuasi-experimental con medidas repetidas. Cada sesión CALL contenía 40 preguntas con diferentes tipos de ayuda cada una.

PREGUNTA NÚMERO ...	NINGUNA RETROALIMENTACIÓN (MAL, INTENTA OTRA VEZ)	DEFINICIONES EN LENGUA META	FRASEOLOGÍA	CONTEXTOS AUTÉNTICOS: CITAS
1	X			
2		X		
3			X	
4	X			X
5		X		
..				
40				X

Tabla 6. Diseño del primer estudio.

#### DATOS CALL

En los análisis de datos CALL, un cero indica que no hubo ninguna respuesta correcta, un 1 significa una respuesta inmediatamente correcta, un 2 apunta a que el estudiante aprovechó una pantalla de ayuda y que ha logrado dar una de las respuestas correctas, mientras que un 3 indica que ha decidido volver a trabajar las palabras que no conocía desde el principio y que dio la respuesta correcta después de ver la clave en la pantalla final (.exit).

El programa activa esta opción después de que los estudiantes completen todas las preguntas y antes de salir definitivamente del programa. Un 3 –más- indica perseverancia, mientras que un cero indica abandono sin volver a la carga (como se ve por la ausencia de treses en el *logfile*, pero un cero también puede apuntar a que el estudiante abandonara sin llegar a la respuesta correcta, incluso después de ver la clave. Muchos treses o puntos más altos indican que se trata de estudiantes tozudos, mientras que los ceros se interpretan por contraste con los treses.

Ahora bien, en esta primera fase del proyecto nos interesan los cambios graduales de las puntuaciones. Si todas las preguntas son igual de difíciles, y si todas las pantallas de ayuda son igual de útiles, en las puntuaciones por pregunta y el promedio de las puntuaciones a lo largo de las sesiones, no habrá variación en cuanto a puntuación por pregunta ni en los promedios. Pero es punto menos que razonable esperar que todas las preguntas o juegos de preguntas sean igual de difíciles. Las diferencias significativas en cuanto a grado de dificultad reflejan las oscilaciones propias de textos naturales.

Los sujetos eran estudiantes de Filología Hispánica que estaban en el segundo semestre de su segundo año de la carrera. El nivel general de estos estudiantes lo describiríamos –de

forma completamente no oficial e informal- como Advanced<sup>1</sup>. Creíamos que su aprendizaje del español se beneficiaría de este tipo de instrucción con ayudas lexicográficas en clases de trabajo individual en laboratorio sin que un profesor dirigiera directamente el proceso de adivinar significados del contexto.

Hemos trabajado con todos los estudiantes. No hemos considerado la posibilidad de constituir dos grupos, por dos motivos. En primer lugar, podríamos recoger gran cantidad de datos a lo largo de cuatro meses que duró el curso, pero los sujetos no abundaban. En segundo lugar, como decíamos antes, el tratamiento formaba parte de un curso normal, por lo cual creíamos que todos los estudiantes deberían poder aprovecharse del experimento / enseñanza. En algunos de los años sucesivos en que hemos realizado el experimento, por motivos de vacaciones que caían en fechas diferentes y las consiguientes variaciones de calendarios, hubo más tiempo para espaciar los tests que en otros, pero los efectos han sido parecidos.

La experiencia colectiva de los alumnos comprendía en todos los casos familiaridad con el uso de ordenadores. Estaban habituados a recibir parte de la instrucción léxica, sintáctica y pragmática mediante laboratorios informáticos. Trabajar con ordenadores era una actividad familiar que no llamaba la atención para nada. Para la mayoría, este programa experimental era simplemente otro programa más. El que más problemas de conciencia tenía era yo, precisamente porque presentar palabras aisladas y sólo después de tener la evidencia de que no sabrían dar ningún equivalente holandés pasar a darles contextos y ayudas, a mí me parecían una manera extremadamente empobrecida e incluso anti-didáctica de obrar. A ellos, nada de esto les parecía extraño. Lo que más me ha sorprendido es ver que se esforzaban cada vez más fanáticamente, apuntando primero todas las palabras de cada capítulo y luego repitiendo toda la serie del capítulo, una y otra vez, hasta sacar la máxima nota posible en el contador automático antes de salir de la sesión.

Las respuestas de cada estudiante (N=18) se almacenan en una de cuatro de categorías: sabe la palabras (puntuación=1), responde después de aviso de error y después de ver la pantalla de ayuda (una de las cuatro definidas para el experimento (puntuación=2), reproduce la respuesta de memoria después de ver la pantalla final (.exit) o de la clave (puntuación=3), o bien abandona en un momento cualquiera o bien abandona después de repetidas tentativas. Un cero indica las preguntas que el alumno no contestó en ningún momento pero también representa las respuestas incorrectas. Es posible que los estudiantes no consiguieran acertar al cabo de varias tentativas, pueden haber abandonado nada más ver la pregunta. El análisis de los conocimientos disponibles indica que no todas las series y subseries de preguntas

---

1 Según los descriptores de escala ACTFL. Ver los párrafos relevantes del texto borrador de los descriptores de producción escrita: <http://www.let.uu.nl/spaans/acc090/guidewrite.htm>, o bien consultar <http://www.actfl.org> y ver las 'guidelines for oral proficiency'.

representaban el mismo grado de dificultad en todas las sesiones.

Las categorías más interesantes son :

- El número de palabras conocidas: si el número de palabras inmediatamente identificadas es demasiado alto, el experimento tendrá demasiadas palabras fáciles y fracasará.
- La proporción de respuestas correctas dadas después de una respuesta inicial incorrecta. Desde luego, la diferencia entre las condiciones de *feedback* / ayuda no desempeña un papel muy importante en la discusión sobre la cantidad de palabras desconocidas; varía el texto original (conocimientos previos, tema, estilo, etc.), pero será evidente que esta variabilidad no nos interesa aquí, ya que nuestra investigación se centra en la proporción de respuestas correctas después de ver la pantalla de ayuda.

La Figura 1 muestra que en cada semana / sesión hay diferencias en número de palabras ya conocidas entre condiciones. Esto se ve por el número de respuestas inmediatamente correctas. En la primera sesión, el porcentaje de palabras conocidas en la condición NF (No Feedback, '*mal, intenta otra vez*') es de un 38%, mientras que en las otras tres condiciones vemos un 30%. En la sesión número 5, las palabras conocidas en la condición NF promediaba un 12%, mientras que en las demás condiciones varía del 18 al 22%. Aunque el porcentaje de palabras conocidas varía entre condiciones y sesiones (como lo indica un análisis de varianza para mediciones repetidas  $F=3.09$ ;  $df=30, 480$ ;  $p < .001$ ), podemos concluir que ninguna de estas sesiones se invalida por desaciertos en la elección de las palabras que había que aprender, o, dicho de otro modo: en ninguna de las sesiones el porcentaje de palabras ya conocidas de antemano es demasiado alto como para no poder estudiar los efectos de las pantallas de ayuda. Teniendo presente el resultado de este primer paso, podemos pasar al siguientes análisis en el que nos centramos en el porcentaje de respuestas correctas, dadas con ayuda de la pantalla de *feedback* / ayuda.

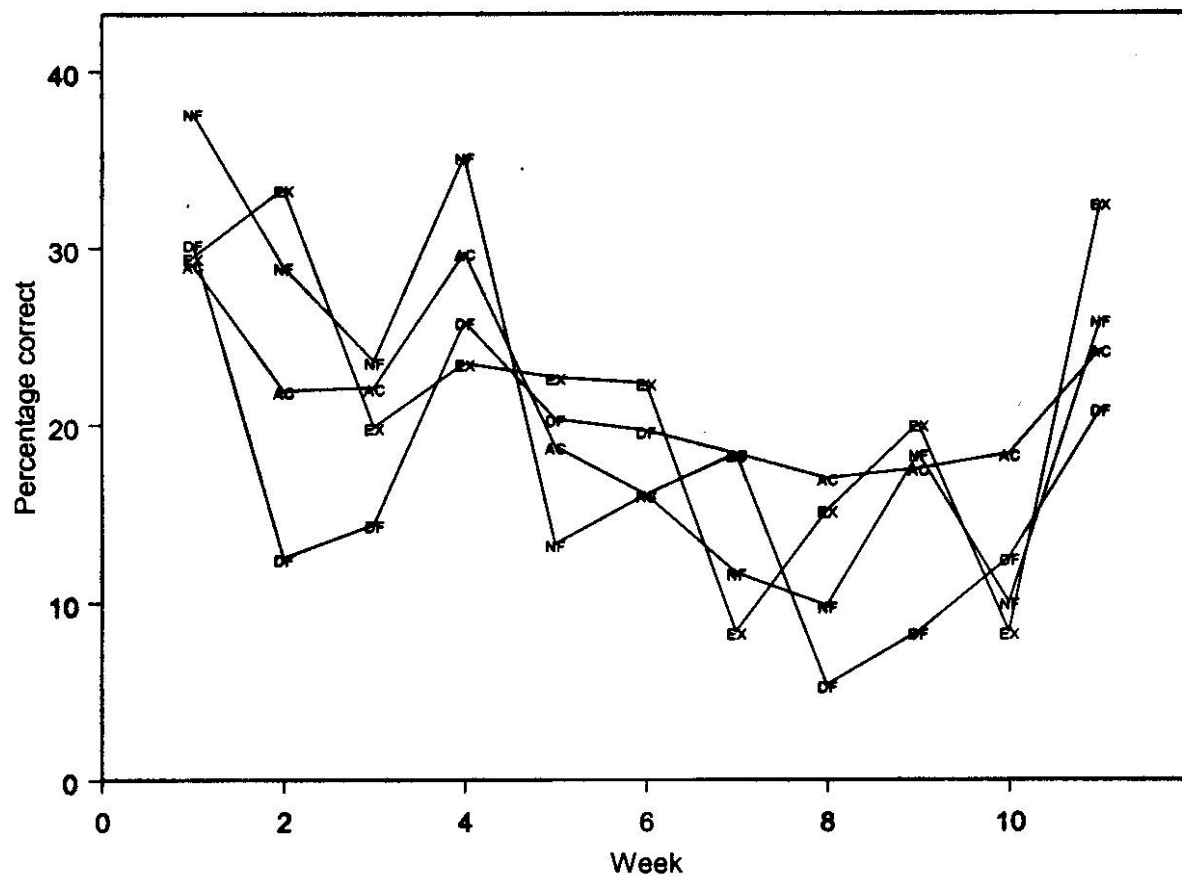


Figura 1. Porcentaje medio de palabras ya conocidas por sesión y por tipo de (NF: No Feedback; DF: DEFinitiones; EX: EJemplos; AC: Contexto Auténtico: citas).

Palabras identificadas después de ver la pantalla con el *feedback*.

En realidad, llevamos a cabo dos análisis: uno sobre el porcentaje observado de respuestas correctas después del *feedback* / ayuda, y otro en el que tenemos en cuenta las diferencias de porcentaje de palabras conocidas (como se ve en la Figura 1).

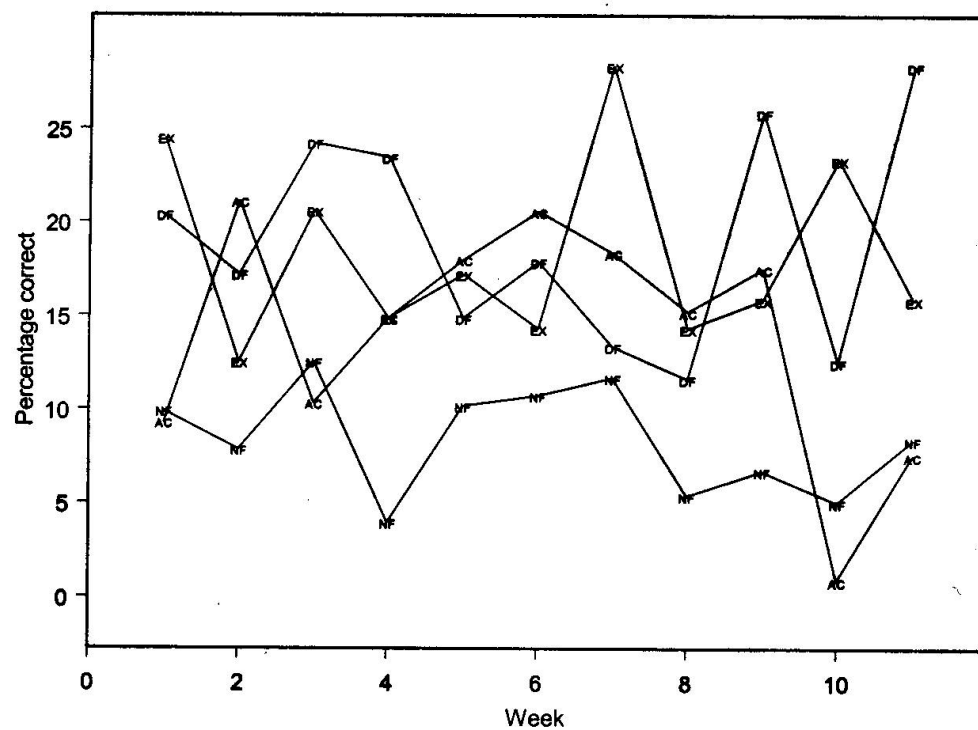


Figura 2. Porcentaje medio de palabras conocidas después del *feedback* por sesión y por condición (NF: No Feedback; DF: DEFinitions; EX: EJemplos, sample sentences; AC: Contexto Auténtico, citas).

En vista de la sorprendente semejanza de los resultados nos adherimos a la primera interpretación que es la más sencilla. La única diferencia radica en el hecho de que el segundo en el que el porcentaje de palabras conocidas se ha usado como covariado: son un poco más marcadas las diferencias entre grupos. En la Tabla 2 se observa el porcentaje medio de respuestas correctas después de ayuda / *feedback*, por tipo de *feedback*, por sesión.

La Figura 2 muestra que en cada una de las sesiones hay diferencias entre las cuatro condiciones de *feedback*. Por ejemplo, en la primera sesión el 25% de las respuestas fue correcto después de un ejemplo o una frase-ejemplo, mientras después de una definición de la palabra en cuestión, un 20% de los estudiantes llegó a dar la respuesta correcta, lo cual contrasta con el 10% de las dos condiciones de 'no *feedback*' y '*contexto auténtico*'.

Las diferencias entre condiciones varía sobre el total de sesiones ( $F=5.69$ ;  $df=30, 480$ ;  $p < .001$ ). También resulta que las diferencias muestran un poco de estabilidad sobre el total de sesiones ( $F=3.80$ ;  $df=10, 160$ ;  $p < .001$ ). Para ser más preciso, en las condiciones '*Ejemplo / frases de fraseología*' y '*Definiciones*', los estudiantes alcanzan el significado correcto más veces que cuando se da la ayuda en términos de contexto auténtico, mientras que la condición de 'no *feedback*' –como era de esperar–, cierra la fila.

Todo junto, podemos concluir que los diferentes tipos de *feedback* dan pie a reacciones diferentes por parte de alumnos extranjeros. Estas diferencias son el reflejo de diferentes tipos de actividad cognitiva. Esto quiere decir que los distintos tipos de *feedback* activan diferentes actividades cognitivas en nuestros alumnos.

#### RESULTADOS EN TÉRMINOS DE GANANCIAS EN CONOCIMIENTOS DE VOCABULARIO

Los resultados evidencian que las series de preguntas contenían una proporción suficiente de palabras desconocidas en cualquiera de las condiciones para permitir que mostráramos que se da un efecto de *feedback*. También hemos podido mostrar que las diferentes formas de *feedback* llevan a diferentes actividades cognitivas. Pero, el diseño experimental en el que los tipos de *feedback* rotaban por cada subserie de cuatro preguntas puede que hayamos subestimado las diferencias entre condiciones. Es muy posible que los alumnos aplicaran sus reacciones ante uno de los tipos de *feedback* a la pregunta siguiente que preveía otro tipo de *feedback*.

Por lo tanto, no parece posible establecer un orden de eficacia entre los cuatro tipos de *feedback*. En retrospectiva, ahora pensamos que los estudiantes aplicaban cualquier tipo de estrategia, indistintamente del tipo de tratamiento o no tratamiento. Son de utilidad pedagógica innegable los procesos de *transfer*, pero en este experimento no dejan de ser



molestos. Tendremos que contrarrestar este tipo de problema si queremos sacar conclusiones sobre la relación entre los conocimientos actuales de los estudiantes y la pantalla de ayuda más provechosa para cada alumno por separado.

Por este motivo, decidimos espaciar los tipos de *feedback* y redactar las preguntas de tal manera que las primeras tres sesiones tuvieran un tipo de *feedback*, la segunda serie de tres tendría otro tipo, y aún otro tipo entraría en acción después de estos dos.

Los tests inicial, intermedio y final dan fe de cambios en la proporción de palabras recordadas gracias al tipo de *feedback* ofrecido. El siguiente párrafo contiene los resultados de estos tres tests.

#### [El pretest, el intermedio y el final de 1997](#)

Como hemos explicado arriba, además de los ficheros CALL que hemos analizado, a los estudiantes (N=18) les dimos un test de 112 preguntas, en tres ocasiones, antes del experimento, un test a mitad de trayecto y otro al final del período. El test B era el reflejo del test A en el sentido que las preguntas de 1-112 se habían ordenado al revés de 112-1. Entre una medición y otra hubo periodos de mes y medio o de dos meses. Las puntuaciones sobre las tres mediciones (pre-, intermedio y post-) varían: son razonables en el primer momento, más bajas en el segundo momento, mientras que la tercera medición, al final del experimento, dio resultados más altos que los de la primera medición. Creemos que estas diferencias en parte se deben al hecho de que las preguntas que formaban parte del tratamiento venían primero, y luego las que no formaban parte del experimento, con lo cual las palabras parcialmente estudiadas constituían la primera mitad del test, y las no estudiadas la segunda parte. Es posible que los estudiantes no hayan tenido tiempo suficiente para leer todas las preguntas. También es posible que las palabras no-tratamiento constituyeran, sin que fuera nuestra intención, un juego de preguntas más difíciles que las otras. Así, puede que los estudiantes no hayan tenido tiempo para contestar correctamente a las preguntas que habían estudiado. El test 3 vuelve a enderezar este posible entuerto repartiendo recompensas y premios por el trabajo realizado en CALL. Los resultados demuestran coherentemente que en términos globales ha habido aprendizaje. Los sujetos contestan correctamente a más palabras tratadas que a las no tratadas. Las puntuaciones globales y las características de las respuestas dan fe de que las preguntas no eran ni demasiado fáciles ni en exceso difíciles.

#### [Segundo estudio \(1998\)](#)

Como explicamos en el párrafo anterior, quedamos satisfechos con las ganancias globales en conocimientos léxicos del 1997, tanto las individuales como las del grupo, ganancias que se basan en la retención de palabras estudiadas y en una creciente capacidad deductiva de los alumnos.

Al mismo tiempo, se imponía afrontar los puntos flojos del diseño de 1997. Queríamos encontrar evidencia más convincente para el uso de los diferentes tipos de *feedback* para medir progresos globales relacionados con los niveles existentes de conocimientos léxicos de los alumnos.

Lógicamente, lo que hacía falta era medir el nivel de conocimientos y analizar progresos en los test de control, para poder elegir el mejor tipo de *feedback* para cada nivel de destreza o cada nivel de conocimientos léxicos.

Había varios motivos para repetir la operación. Habíamos podido demostrar que la forma de la tarea influye en la ejecución de la misma, pero no habíamos podido demostrar efectos de cada uno de los tipos de *feedback*. En segundo lugar, el número de estudiantes en los cursos de los que nos servíamos era reducido.

En el segundo estudio, la pregunta central es la que gira en torno a los diferentes grados de eficacia de los varios tipos de *feedback* durante el aprendizaje léxico. Sin embargo, queríamos evitar los puntos débiles del primer estudio (ver arriba). Por eso hemos cambiado ligeramente el diseño del estudio (Tabla 7). En lugar de ofrecer todos los tipos de *feedback* en casa sesión, nos limitamos a un solo tipo. De hecho, optamos por trabajar con un mismo tipo de *feedback* en series de tres sesiones consecutivas. Al cabo de cada serie de sesiones medíamos los conocimientos léxicos de los estudiantes (N=11-13) con preguntas de entres las palabras tratadas en las tres sesiones precedentes del experimento CALL/ELAO, al igual que con preguntas basadas en palabras que habíamos excluido del experimento.

Tal y como se ve en la Tabla 7, hemos medido los conocimientos de 23 palabras que habíamos presentado sin *feedback* ('*mal, intenta otra vez*') tanto al comienzo como al final de la primera serie de sesiones. Al mismo tiempo combinábamos estas 23 preguntas con 56 otras en esos dos momentos de medición. En cuanto a la condición de *feedback* mediante definiciones, hemos medido la familiaridad con 19 palabras, y en tres ocasiones: al comienzo del estudio, al final de la primera serie de sesiones, y al final de la segunda serie de sesiones. Las 13 preguntas características de la condición de "ejemplo" han sido medidas en cuatro ocasiones, al igual que las palabras "nuevas".

Con este diseño, la condición de definiciones y la de ejemplos sirven como condición de control para la de '*mal, intenta otra vez*'. En vista de que estas condiciones no se enseñan en la primera serie de sesiones, no podemos esperar ganancias sustanciosas para esas preguntas, y no podemos esperar nada que ganancias de menor importancia para '*mal, intenta otra vez*'. En cuanto a la condición de definiciones, las ganancias pueden compararse con los resultados de las mismas preguntas de la primera parte del curso, pero también con las ganancias de la

condición de ejemplos. Para esta última condición sólo podemos comparar con las mediciones anteriores.

	PREGUNTAS			PREGUNTAS			PREGUNTAS			PREGUNTAS	
	E	N	Sesiones	E	N	Sesiones	E	N	Sesiones	E	N
No feedback	23	56	1-3	23	56						
Definiciones	19			19		4-7	19	56			
Ejemplos	13			13			13		8-11	13	56

Tabla 7. Diseño (E: palabras estudiadas en CALL/ELAO; N: palabras Nuevas que NO formaban parte del tratamiento en CALL/ELAO).

#### Resumen de los resultados de 1998

De la Figura 3 se desprende que los estudiantes no muestran ganancias en la condición de 'mal, intenta otra vez'. La diferencia entre el test inicial y el final (primero y segundo, respectivamente) no es significativa ( $t = .76$ ;  $df = 11$ ;  $p = .47$ ). Tampoco vemos diferencia alguna entre las tres condiciones en el momento de la primera medición.

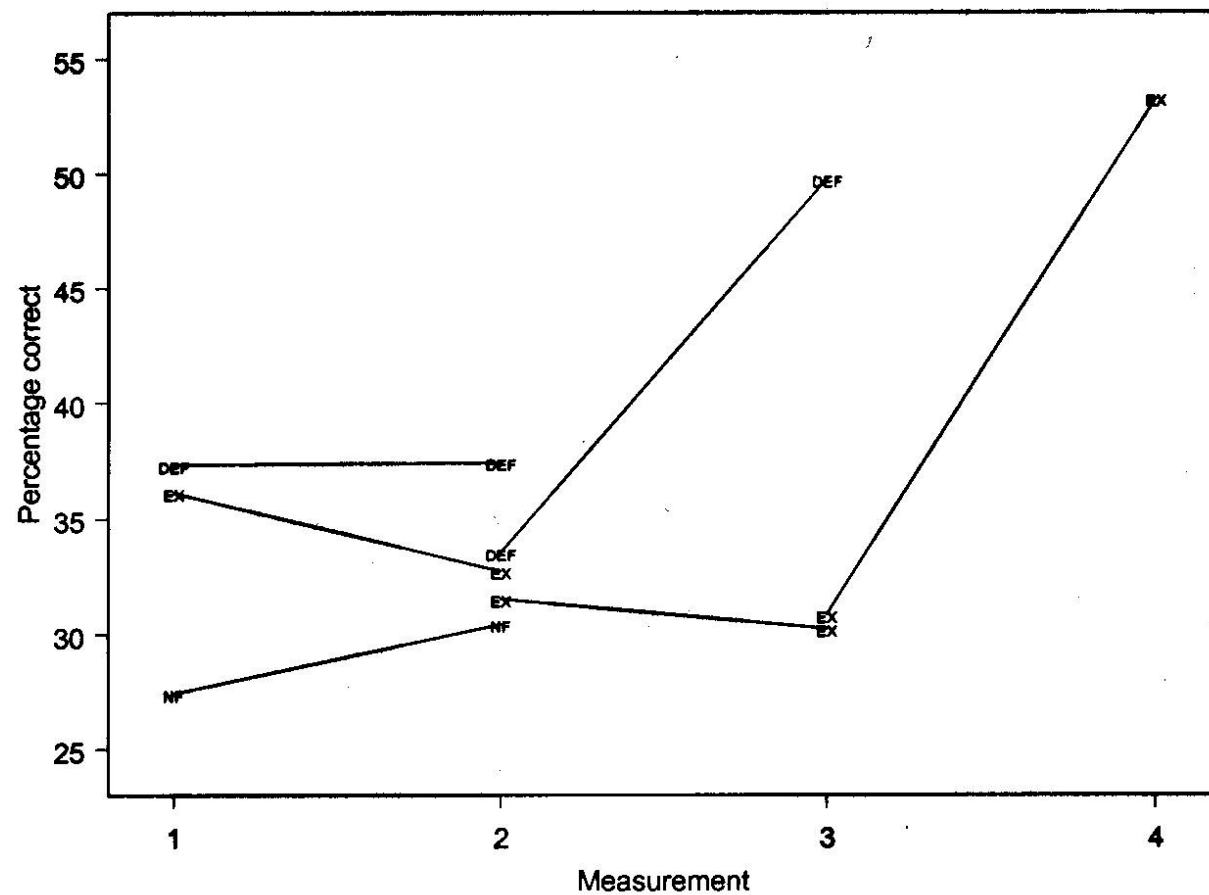


Figura 3. Resumen de los resultados de 1998.

El *feedback* ofrecido en términos de definiciones aumenta los conocimientos. Los conocimientos léxicos de los estudiantes mediante el uso de definiciones aumentan entre el segundo y tercer momento de medición ( $t=3.11$ ;  $df=10$ ;  $p=.01$ ), mientras que no hay diferencia entre estas preguntas en la primera y la segunda medición ( $t=.01$ ;  $df=11$ ;  $p=1.00$ ), como tampoco la hay entre la segunda y la tercera medición para la condición de ejemplos ( $t=-.65$ ;  $df=11$ ;  $p=.53$ ).

Igualmente, las preguntas del tipo de frases ejemplo tampoco arrojaron diferencias en conocimientos léxicos entre el segundo y tercer momento de medición ( $t=-.33$ ;  $df=10$ ;  $p=.75$ ), pero en el último período se observan ganancias significativas ( $t= .51$ ;  $df=11$ ;  $p=0.005$ ).

Podemos concluir, por tanto, que los dos tipos de *feedback*, mediante definiciones y mediante frases ejemplo, son más rentables que el *feedback* del tipo '*mal, intenta otra vez*'.

#### Transfer

La segunda pregunta que nos gustaría contestar concierne el *transfer* de estrategias usadas para llegar al significado de palabras nuevas que no formaban parte del tratamiento. Los resultados se recogen en la Tabla 8.

MEDICIONES				
N	1	2	3	4
13	22.0 (11.9)	36.1 (11.4) <sup>a</sup>		
11		36.8 (12.5)	50.6 (9.4) <sup>b</sup>	
11			53.2 (10.2)	60.9 (14.2) <sup>c</sup>

Tabla 8. Promedios (porcentajes) y desviaciones estándar (entre paréntesis) para palabras nuevas en diferentes ocasiones (N = número de estudiantes), a:  $t= .87$ ;  $df=11$ ;  $p=.002$ ; b:  $t=3.73$ ;  $df=10$ ;  $p=.004$ ; c:  $t=2.26$ ;  $df=10$ ;  $p=.047$ .

En la primera medición, la puntuación media de los 13 estudiantes que se presentaron al test que incluía nuevas palabras ascendía a un 22%, igual que en el segundo momento. Después de recibir NF ('*mal, intenta otra vez*') su puntuación media subió a un 36.1%. Después del segundo tipo de tratamiento (definiciones), su puntuación media aumentó en 14 puntos, y

después del tratamiento a base de frases ejemplos volvemos a ver otro aumento, esta vez de un 17%. Las ganancias en conocimientos léxicos son significativas para las comparaciones.

#### Conclusiones en torno a los test inicial e intermedio y el post-test de 1998

Concluimos que las puntuaciones en las preguntas que no habían entrado en el material de tratamiento experimental crecen en todos los momentos de medición. Esto significa que ha habido *transfer* en todas las etapas del experimento. Las ganancias más altas se han dado después de trabajar con las pantallas de ayuda que contenían definiciones. Pero incluso cuando los estudiantes y nosotros nos acercábamos al final del curso, y a la vista de las ansiadas vacaciones de verano, las puntuaciones entre el punto 3 y 4 seguían aumentando después de trabajar con las frases ejemplo. Pese a disponer de grandes cantidades de datos, no cabe duda de que el número de sujetos seguía siendo bajo. Por eso decidimos repetir el experimento durante el siguiente año académico, con una nueva y fresca generación de estudiantes.

#### Tercer estudio (1999)

El tercer estudio se sirvió del mismo diseño que el segundo. Los resultados son semejantes a los del año 1998. Entre la primera y la segunda medición no se evidencia aumento en conocimientos léxicos en ninguna de las tres ( $p > .47$ ). Entre la segunda y la tercera medición sólo hay ganancias. Aumenta la condición de definiciones ( $t=2.83$ ;  $df=6$ ;  $p=.03$ ), mientras que entre la tercera y la segunda ocasión se produce una clara ganancia para la condición de ( $t=2.70$ ;  $df=6$ ;  $p=.035$ ).

La segunda pregunta que hemos formulado en el estudio de 1999 se centraba en la relación entre conocimientos de vocabulario y la eficacia de la ayuda ofrecida en las tres. Es posible que el aumento en conocimientos léxicos no dependa sólo de la condición sino también del nivel léxico de los estudiantes. Los resultados de este análisis se hallan en la Figura 4.

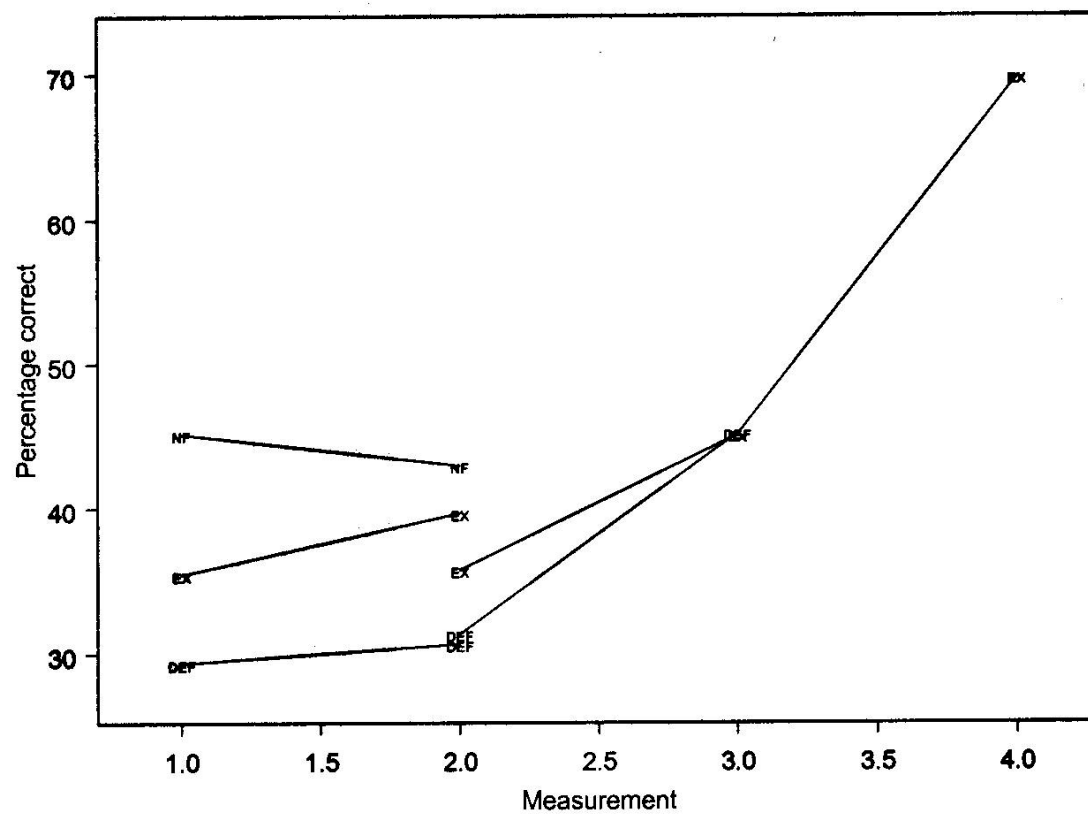


Figura 4. Representación gráfica de las puntuaciones medias por condición y por medición (en porcentajes; NF: No Feedback; DF: DeFiniciones; EX: Ejemplos).

### La relación entre el feedback y conocimientos

La Figura 5 muestra que la relación entre conocimiento léxicos (el eje x) y el porcentaje de palabras correctas (el eje y) es más fuerte en la condición de definiciones y de ejemplos que en la condición de '*mal, intenta otra vez*'. Esto quiere decir que la mayor parte de los estudiantes con más conocimientos léxicos que sus compañeros se benefician más de las definiciones y los ejemplos. Para los estudiantes con conocimientos léxicos por debajo de un 15%, la condición de '*mal, intenta otra vez*', en cambio, es la que da las puntuaciones más altas. Para aquellos estudiantes que se encuentran entre el 15% (aproximadamente) y el 27% en conocimientos léxicos, las definiciones en lengua meta parecen ser las más adecuadas y el mejor tipo de ayuda, mientras que para los que demuestran un porcentaje más elevado de palabras conocidas, ejemplos en lengua meta parecen proporcionar el *feedback* más eficaz.

Obsérvese que los resultados (los porcentajes) dependen bastante del instrumento. Lo que de verdad necesitábamos era un test estandarizado de conocimientos léxicos. A lo largo de los años que ha durado el experimento nos hemos servido de este test de fabricación casera, pero de todos modos ha quedado claro que la condición NF (No Feedback, '*mal, intenta otra vez*') es la que menos ganancias en puntuación arroja. EX (frases ejemplo) empieza muy bajo. Su eficacia para alumnos con bajos conocimientos iniciales es muy baja al comienzo, pero aumenta rápidamente a lo largo del curso. Las DEF (definiciones en lengua meta) contribuyen más que la condición NF, pero menos que la fraseología, y siguen subiendo en una línea bastante marcada y ascendente hacia la puntuación del test intermedio. En el test final, las DEF contribuyen considerablemente, pero siempre menos que EX, las frases ejemplo.



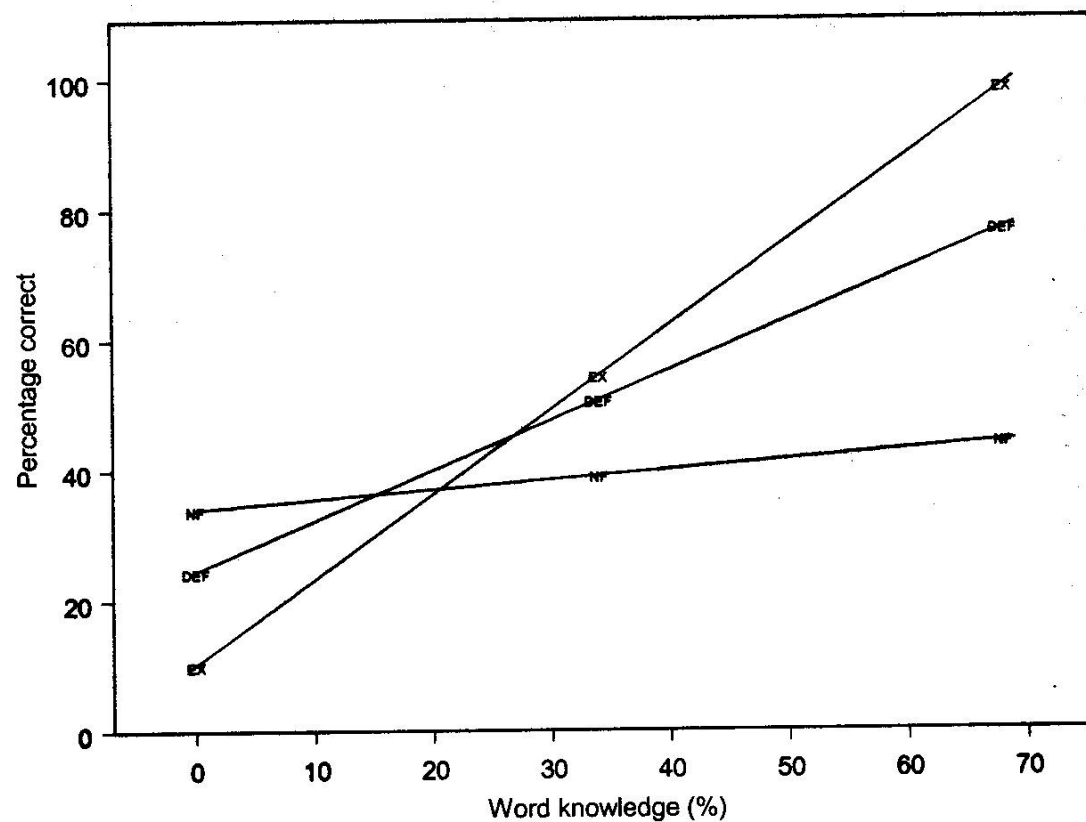


Figura 5. Relaciones extrapoladas entre conocimientos de palabras y porcentajes de respuestas correctas, por condición (NF=No Feedback (mal, intenta otra vez'), DEF = DEFinición en lengua meta), EX = EJemplo (fraseología)

Concluimos que para estos alumnos intermedio / avanzados la condición NF no estimula el acceso a conocimientos disponibles pero no usados. Esto es positivo porque si fuera al revés este programa se caracterizaría más como test que como programa de aprendizaje de nuevas estrategias cognitivas. Para el conjunto de los alumnos, los ejemplos y las definiciones parecen funcionar igual de bien, o igual de mal durante las primeras sesiones, mientras que los ejemplos que usan fraseología dejan atrás los resultados de las definiciones a partir de la segunda mitad de las sesiones.

Puede explicarse esto subrayando el hecho de que frases tomadas de la parte de fraseología –a condición de estar a la altura de razonables exigencias lexicográficas– ofrecen información implícita en torno a todos los puntos que detalla Richards (1985: 183). Estos datos también revelan que las definiciones pueden funcionar mejor en niveles más bajos. Una definición apunta un solo significado mientras que la fraseología requiere mecanismos de procesamiento más elaborados que parecen aprendibles y que pueden reforzarse a lo largo del curso. También puede que sean más complejos y difíciles de aprovechar. Esto pudiera explicar el hecho de que sólo después de que pase cierto tiempo los alumnos intermedios/avanzados parecen beneficiarse de este tipo de ayuda.

### Discusión

Los resultados de este estudio –especialmente los datos recogidos en 1997– indican que el tipo de *feedback* ('mal, intenta otra vez', definiciones redactadas en la lengua meta, fraseología y citas de textos auténticos) influye en la conducta del alumno. El tipo de *feedback* también incide en la retención a largo plazo. En los datos recogidos en 1998 y 1999, hemos variado el tipo de *feedback* de acuerdo con la puntuación de los estudiantes en los tests inicial, intermedio y final, y hemos podido comprobar un efecto en la retención. El dato más importante, sin embargo, es que la noción de qué constituye *feedback* eficaz depende de los conocimientos disponibles en cada momento. Este resultado lo interpretamos como señal de que los materiales diseñados para ELAO, o para ICT, debería partir del nivel "contemporáneo" del alumno individual, en vez de tratar a todos los alumnos como si tuviesen un solo nivel homogéneo.

En este estudio me he servido de trabajos anteriores en el campo de maneras de dar *feedback* y en el de aprendizaje de vocabulario. He recurrido a la información disponible en un extenso diccionario bilingüe editado en forma "tradicional", y cuidadosamente copiada al formato ELAO. Creo que en un futuro no demasiado lejano, si esta cuestión interesa a editoriales lexicográficas de diversos países que entre ellas producen productos diversos, destinados, sin embargo, al mismo tipo de público, podríamos pensar en un diccionario del alumno en línea, flexible en el sentido de que se ajuste al nivel de destreza y conocimientos y

que pueda acompañarle en su viaje hacia la maestría poniendo cada vez más información a su disposición.

No puedo dejar de terminar con unas notas realistas. En primer lugar, soy consciente del hecho de que las editoriales suelen querer vender libros, no abonos a internet. En segundo lugar, no todo funciona de la misma manera en todo momento. Los estudiantes cambian, el mundo cambia y los profesores también. Pero ahora que hemos llegado a este punto, quizás el momento apropiado para montar no sólo este diccionario futurista sino también crear métodos para analizar, describir y medir los efectos de tantos entornos virtuales que hay ya en el mundo, y de cuya eficacia en realidad aún sabemos muy poco.

## EPÍLOGO

Desde la lectura de mi tesis han pasado ahora un par de años y ha habido momentos en que la idea de crear un diccionario del alumno amoldable al nivel que éste pueda tener en diferentes momentos de su trayectoria me ha parecido cosa de locos. Que yo sepa, no existen diccionarios que combinen información típica de los monolingües con otra no menos valiosa de los bilingües. Unir esfuerzos y crearlos sería un paso gigantesco por parte de las editoriales especializadas. Ciertamente existen cada vez más versiones de diccionarios en CD-ROM, de infinita facilidad de uso para quien no se separa nunca de su ordenador, para quien no lo apague nada más que para irse a la cama a dormir de vez en cuando. Para los auténticos adictos y los que quieren tener algo bienoliente entre manos están las versiones tradicionales de papel. Ciertamente es también que existen cada vez más sistemas de medición de nivel de destrezas lingüísticas. El Instituto Cervantes tiene un test de auto-evaluación, existen varios en los Estados Unidos y sin duda alguna ya varios más que no he podido conocer aún, y habrá muchos más en el futuro. Ciertamente que hubo sistemas electrónicos, utilizables en laboratorios informáticos que permitían registrar todas las consultas que hiciera un alumno a un diccionario electrónico instalado en el servidor de ese mismo laboratorio: para saber lo que hacía, con qué tesón se metía en la información disponible y cuál pudiera ser el resultado de la calidad y profundidad de sus consultas sobre el trabajo (lectura, redacción, traducción) que estuviera realizando en ese momento. Tuvimos un programa electrónico en nuestros laboratorios informáticos de la Universidad de Utrecht que han funcionado hasta tiempos recientes. Sandor Spruit (1995) echó las bases para un sistema de ese tipo y Elena Landone (2000) describe un experimento que bien pudiera realizarse, aún hoy.

Lo que falta a todas luces es combinar estos tres componentes. Ése es mi sueño, y ahora al volver sobre mis pasos, releendo y adaptando las páginas del capítulo 7 de mi tesis, creo que es un sueño con futuro, no que todos tengamos este sueño per saecula saeculorum, sino del tipo que alguien, en algún momento, volverá a hacerlo suyo para convertirlo en un valioso

instrumento de enseñanza, y de investigación.

## BIBLIOGRAFÍA

Alvar Ezquerro, M. (1995). *Diccionario para la enseñanza de la lengua española*. Vox / Alcalá de Henares, Barcelona: Bibliograf.

Juilland, A.G., & Chang-Rodríguez, E. (1964). *Frequency dictionary of Spanish words*. The Hague: Mouton.

Landone, E. (2000). Diccionarios electrónicos: nuevos objetos culturales con implicaciones en la metodología de la investigación y en la didáctica de E/LE. *Didattica della lingua spagnola. Recherche*. 3. CUEM. *Enseñanza del léxico. Talleres didácticos de 1998/1999*. Edición al cuidado de M. Scaramuzza Vidoni, pp. 15-41.

Richards, J.C. (1985). *The Context of Language Teaching*. Cambridge: Cambridge University Press.

Schachter, J. (1991). Corrective feedback in historical perspective. *Second Language Research*, 7, 89-102.

Slagter, P.J. (1992). *Van Dale Handwoordenboeken Spaans-Nederlands en Nederlands-Spaans*. Utrecht: Van Dale Lexicografie.

Slagter, P.J. (2000). *Learning by instructing. Studies in Instructed Foreign Language Teaching*. Tesis doctoral, Universidad de Utrecht, Países Bajos.

Spruit, A.G.L. (1995). *An electronic bilingual dictionary for educational use*. Universiteit Twente, Master's thesis in ICT.