

Moreno Teva, Inmaculada (2012), *Las secuencias formulaicas en la adquisición de español L2*. Tesis doctoral. Universidad de Estocolmo, 225 págs. ISBN 9789174474985

La investigación *Las secuencias formulaicas en la adquisición de español L2*, de Inmaculada Moreno Teva, publicada en mayo de 2012 como tesis doctoral en el Departamento de Español, Portugués y Estudios Latinoamericanos de la Universidad de Estocolmo, aborda uno de los aspectos más complejos del proceso de adquisición-aprendizaje de una lengua (tanto materna como extranjera): las secuencias lingüísticas preestablecidas, esto es, las combinaciones de palabras fijadas por el hábito o costumbre que prefieren los hablantes de una lengua. Implícitamente, pues, el trabajo busca dar respuesta a la compleja pregunta sobre lo que significa dominar una lengua: he aquí la motivación principal. Se parte del supuesto de que para aprender un idioma no solo es necesario conocer sus palabras individuales, sino también saber cómo se combinan adecuada e idiomáticamente.

Por su carácter convencional, emplear secuencias formulaicas (“prefabricadas”) en una lengua otorga fluidez al hablante de un idioma. Estas combinaciones estables de palabras en el trabajo que reseñamos son definidas por la autora como “unidades léxicas que constan de dos o más palabras, secuencias de palabras que se repiten de una forma determinada, han sido memorizadas y están sometidas a un menor o mayor grado de fijación” (Moreno Teva 2012: 15). Bajo esta definición, se analizan en esta investigación doctoral secuencias de palabras de muy distinta naturaleza: combinaciones con significado léxico –*hacer la pelota, mercado laboral, a largo plazo, fácil de usar, ser un rollo*–, combinaciones gramaticales –del tipo *todo el mundo*, locuciones verbales de obligación como *tener que* o *deber de*, marcadores del discurso como *en cambio, no obstante, por eso*–, esquemas léxico-sintácticos recurrentes –*lo interesante es que...*– y rutinas conversacionales –*buenos días* o *muchas gracias*, por ejemplo–.

El objetivo principal de la tesis es analizar cómo evoluciona el uso de estas secuencias de palabras en hablantes suecos de español L2 durante un periodo de estancia en España, de unos cuatro meses aproximadamente, para comparar este uso y su evolución con el empleo de estas mismas secuencias formulaicas por parte de hablantes nativos de español. La tesis persigue además otro objetivo secundario, que es el de estudiar el efecto que causa en el uso de estas combinaciones de palabras el tipo de actividad comunicativa en la que participan los hablantes suecos de español L2 estudiados (y también los hablantes nativos), concretamente en dos tipos de actividades: la *negociación* y el *grupo de discusión*, interacciones conversacionales en ambos casos.

Para lograr estos dos objetivos –a) analizar el efecto de la estancia en el extranjero; y b) estudiar el efecto del tipo de actividad comunicativa en el uso de secuencias formulaicas–, la tesis recoge los datos que proporciona un corpus de español oral ya constituido previamente en el Departamento de Español, Portugués y Estudios Latinoamericanos de la Universidad de Estocolmo: el corpus

AKSAM: *Competencia discursiva y sociocultural en hablantes nativos y no nativos de español*. El corpus AKSAM está formado por conversaciones dirigidas o semidirigidas, que mantienen, por un lado, hablantes no nativos con hablantes nativos y, por otro, hablantes nativos entre sí. Así pues, recoge conversaciones de hablantes nativos exolingües y endolingües. Al disponer de datos, por un lado, de aprendices de español y, por otro lado, de hablantes nativos, se comparan en el trabajo los usos de expresiones formulaicas en ambos tipos de hablantes tanto en cuanto a cantidad como en cuanto a variación<sup>1</sup>.

Un primer escollo que tiene que esquivar esta investigación es el de la propia unidad de análisis: las llamadas en el trabajo *secuencias formulaicas* y sus tipos. Cómo distinguir una combinación de palabras formulaica de otra que no lo es, y cuántos tipos pueden identificarse no es tarea fácil. Destaca la numerosa bibliografía existente sobre combinaciones estables de palabras, tanto para el análisis de otras lenguas como para el español: en la tesis que reseñamos se recogen todos estos trabajos. A pesar de ello, no existe mucho acuerdo sobre el lenguaje formulaico ni en su definición ni en los criterios de identificación; tampoco hay consenso sobre los tipos de expresiones formulaicas que considerar. Las combinaciones estables de palabras han sido abordadas en el plano del texto-discurso, en que interesan como *fórmulas* que ayudan a construir el discurso (Coseriu 1992); otra perspectiva relevante es la que analiza estas expresiones como unidades del plano léxico, como *unidades léxicas polixemáticas* (Corpas Pastor 1997) que se han de aprender en bloques, “semiconstruidos”, según la denominación que adopta el *Plan curricular del Instituto Cervantes* (2006); otro ángulo de aproximación es el del enfoque léxico (cf. Lewis 1993), en el que se concibe la lengua como *léxico gramaticalizado*. En este sentido, se echa en falta en la tesis que reseñamos en qué plano o nivel de análisis se analizan las secuencias formulaicas objeto de estudio.

Ante la falta de acuerdo, en la tesis se opta por tener en cuenta varios criterios para identificar una secuencia formulaica: se distinguen en el trabajo criterios, por un lado, “necesarios”, y criterios, por otro lado, “típicos”. Los necesarios son cuatro: secuencia de como mínimo dos palabras o morfemas libres, coherencia fonológica, variabilidad total o parcialmente limitada y unidad de sentido; los típicos son ocho: frecuencia de uso, asociación a un contexto determinado, vinculación a uno o varios usos pragmáticos específicos, añadido semántico (significado diferente y más restringido de aquel que sugieren sus componentes), mayor complejidad, reflejo de un uso inapropiado como consecuencia de su

---

<sup>1</sup> Las preguntas de investigación que se plantea la tesis son las siguientes: ¿qué diferencias generales existen entre nativos y no nativos en su uso de las secuencias formulaicas?; ¿qué efecto produce en las secuencias formulaicas de los hablantes no nativos la estancia durante un periodo de tiempo prolongado en un país de habla hispana?; ¿en qué parámetros y en qué medida aparece una significativa variación individual en las secuencias formulaicas de los hablantes no nativos? ¿Se producen cambios en la variación individual relacionados con los cambios en la cantidad y la variedad de las secuencias formulaicas?; ¿repercute el tipo de actividad realizada por los hablantes no nativos en los resultados obtenidos? Si es así, ¿de qué modo?

memorización holística, uso idiosincrásico, y posibilidad de cambio lingüístico (cf. Moreno Teva 2012:73).

No obstante la cantidad de criterios que se detallan para identificar las secuencias formulaicas objeto de estudio, no queda claro en la tesis el peso que se otorga a cada uno de ellos. Además, por una parte, algunos de los criterios (como el relativo al reflejo de un uso inapropiado) solo son aplicables a un tipo determinado de secuencias formulaicas; y, por otra, no se aplican con coherencia algunos de ellos, como es el caso de la frecuencia de uso: se consideran en el trabajo secuencias formulaicas combinaciones de palabras cuya frecuencia de uso en el corpus analizado es baja. Expresiones como *venir a SN*, *estar en*, *pasarse a o el año que viene* son analizadas como secuencias formulaicas en la investigación, a pesar de su baja frecuencia de uso en el corpus objeto de estudio (son estructuras que presentan 4 ocurrencias en los datos analizados).

En cambio, no se recoge un criterio que consideramos relevante para calibrar la idiomatidad de una expresión: su grado de lexicalización. Una expresión recogida en el trabajo como *hacer la pelota* se usa en un sentido metafórico, traslaticio en español en contextos informales; en cambio, una combinación como *estallar la guerra*, también analizada en la tesis, se emplea en su sentido literal; no obstante, en el análisis se tratan ambas combinaciones del mismo modo. Pero *hacer la pelota* ha experimentado un proceso de lexicalización y, por lo tanto, constituye una unidad léxica; por su parte, el sintagma *estallar la guerra* no está lexicalizado, sino que está constituido por dos palabras que suelen combinarse con frecuencia pero que no forman una única pieza léxica. En el ejemplo de *hacer la pelota* estamos ante una unidad verbal permutable por un verbo como *adular*, *alabar* o *lisonjear*; en el ejemplo de *estallar la guerra* no es posible este tipo de sustitución por una sola unidad verbal: se trata, pues, de un sintagma formado por dos unidades léxicas. Y esta diferencia es debida a que son combinaciones de palabras de distinta naturaleza, que en el trabajo que reseñamos se abordan como secuencias formulaicas del mismo tipo.

El marco teórico del trabajo tiene en cuenta sobre todo la propuesta pragmático-comunicativa seguida por Wray (2002), quien defiende la idea de un continuo en el grado de fijación de las secuencias de palabras, desde secuencias muy fijadas, con todas sus formas léxicas codificadas (como los refranes, los dichos, las rutinas conversacionales) hasta otras más abiertas, con cupos abiertos (*open slots*), en que pueden darse variantes en las combinaciones según las necesidades individuales de cada participante en la interacción, como, por ejemplo, en la estructura *venir a SN*. En este último caso las secuencias de palabras no tienen elementos léxicos fijos, pero sí especificados semántica o gramaticalmente. Este continuo diluye la distinción entre léxico y gramática, y establece la siguiente gradación en la tipología de combinaciones de palabras: *unidades fijas* > *unidades semifijas* > *marcos abiertos*.

Desde nuestro punto de vista, con este planteamiento se superponen en el trabajo el análisis de unidades léxicas con el de estructuras sintácticas, sin especificar que estamos ante planos distintos. Esto es, no se parte de una base

común de comparación en los tipos de expresiones que se estudian: en unos casos es la unidad léxica, esto es, la palabra la que constituye el eje en torno al cual se organiza la secuencia formulaica; en otros casos, la secuencia formulaica no es estrictamente una combinación de dos o más palabras sino un esquema sintáctico. Se requeriría aclarar si la unidad de base para el análisis es la palabra, como se infiere de la definición de secuencia formulaica que se ofrece, o la construcción sintáctica. Esta aclaración permitiría revisar el hecho de que en el trabajo se consideren fórmulas construcciones como *la gente* o *un poquito*, cuya función en la interacción se explicaría mejor desde un punto de vista morfosintáctico más que formulaico.

Esta distinción es también pertinente en relación con la adquisición de una lengua. Interesa en la tesis observar cómo se adquieren estas unidades lingüísticas en el proceso de aprendizaje de una L2. Según el modelo de Wray (2002 y 2008) que se sigue, las secuencias de palabras son tratadas holísticamente en general por un hablante nativo, no son descompuestas a no ser que sea necesario y siempre lo menos posible. En el continuo de fijación léxica y semántica, la producción “holística”, en secuencias formulaicas, es la forma como nos comunicamos; el procesamiento “analítico” se lleva a cabo según la situación de comunicación, sostiene la autora de la tesis. No obstante, en el proceso de adquisición-aprendizaje de una segunda lengua cabe considerar el procesamiento holístico (por ejemplo, de *buenas tardes*) como un primer estadio de aprendizaje; a medida que un aprendiz va avanzando en su dominio del idioma que aprende empieza a descomponer estas combinaciones, a analizarlas (*buenas*-adjetivo femenino plural y *tardes*-sustantivo femenino plural) y jugar con ellas (como hace un nativo: *malas tardes*); esta variación en el uso y toma de conciencia de los componentes de una expresión es un indicio de progresión en el nivel de dominio de una L2.

Son varias las razones que se aducen en el trabajo en relación con el procesamiento holístico de las secuencias formulaicas que se defiende en la tesis: un hablante procesa holísticamente las combinaciones de palabras fundamentalmente para aliviar cargas de análisis lingüístico (si una secuencia regular de palabras es retenida como una unidad léxica, se evita el procesamiento analítico), para estructurar el discurso y guiar al oyente o para señalar identidad y pertenencia a una comunidad de habla. Wray postula que todas estas funciones están supeditadas a un objetivo principal: la promoción de los intereses del hablante, esto es, el logro de su éxito en la interacción. Desde otra perspectiva, estas formas de hablar prototípicas garantizan la comprensión de nuestro interlocutor. Podrían considerarse, pues, a diferencia de como se plantea en el trabajo, un recurso que garantiza el acercamiento a nuestro destinatario en la interacción, al usar combinaciones que le resultan familiares, convencionales. Desde nuestro punto de vista, las expresiones idiomáticas permiten ser más cooperativos comunicativamente, y compartir hábitos (lingüísticos).

En cuanto a los tipos de secuencias formulaicas, en el trabajo que reseñamos se establece una primera distinción entre secuencias formulaicas *idiomáticas*, compartidas por los hablantes nativos, secuencias formulaicas *no idiomáticas*, no

percibidas como idiomáticas por los nativos aunque funcionan así para los no nativos (como parte de su interlengua), y secuencias formulaicas *fallidas*, carentes del sentido y la función que persiguen. No es clara esta clasificación en el trabajo, pues no se discuten ejemplos; además, no son transparentes los criterios que se consideran para señalar de qué tipo son; para las secuencias *fallidas*, por ejemplo, se establecen los siguientes criterios: son traducciones literales de otros idiomas sin correspondencia en español, se usan en contexto inadecuado o contienen errores léxicos que afectan al sentido, los cambios de código lingüístico que correspondan a una secuencia formulaica [sic] (Moreno Teva 2012: 85); es difícil comprender estos criterios sin ilustrarlos con ejemplos analizados. En el capítulo que se dedica al estudio de los datos del corpus se comentan algunos ejemplos, pocos; sería necesario establecer sus diferencias con mayor claridad cuando se presenta esta tipología de secuencias formulaicas.

A partir de esta primera clasificación, siguiendo la propuesta de Erman y Warren (2000), se caracterizan tres tipos de secuencias en un segundo nivel de análisis: un primer tipo de secuencias formulaicas conceptuales (*mercado laboral, llegar a un acuerdo, consultar a, a pies juntillas, de buen humor*), que pueden ser nominales, preposicionales, adjetivales, adverbiales, verbales o proposicionales (del tipo *hola, ¿qué tal?*; o secuencias frásticas como *hace diez años que no le veo*), un segundo tipo de secuencias formulaicas gramaticales (*todo el mundo, algún que otro*), que a su vez se clasifican en pronominales, modales, temporales y aspectuales, negaciones, cuantificadoras o de gradación; y un tercer tipo de secuencias formulaicas discursivas (*por eso, lo que está claro es que*), que cumplen un papel externo más allá de la función predicativa, con varios subtipos: organizadoras del discurso, interregulativas, autorregulativas y expresiones aproximadoras.

Hubiera sido muy útil contar con un apéndice en el trabajo que reuniera todas las unidades léxicas analizadas en la investigación clasificadas según esta tipología. Resulta difícil establecer una lista concreta de elementos dentro de cada categoría; de todos modos, en la discusión de los resultados se muestran algunos ejemplos, pero no todos los que se han considerado dentro de un tipo particular. En un trabajo doctoral de estas características es necesario presentar todas las secuencias formulaicas identificadas en el corpus reunidas en un apéndice, para su consulta.

Se realiza además un análisis cuantitativo, que ha consistido en contabilizar cuántas ocurrencias y tipos de secuencias formulaicas se dan por 100 palabras gráficas en cada hablante nativo y no nativo del corpus analizado. Se recogen en el trabajo las frecuencias absolutas de aparición de las combinaciones de palabras, pero habría sido más claro considerar su dispersión entre los hablantes analizados: es posible que un hablante recurra siempre a una misma secuencia y, en cambio, esta no aparezca nunca en los otros participantes de la conversación. Por lo tanto, la frecuencia de uso no solo tiene que ver con las frecuencias absolutas, sino también con el número de interlocutores que la emplean: a mayor número de hablantes que recurren a una misma combinación de palabras mayor grado de

convención presentará.

Para el estudio evolutivo que se ha querido llevar a cabo en el trabajo sería necesario completar el corpus de datos utilizado. Como hemos dicho, uno de los objetivos de la tesis es analizar la variable estancia en el extranjero para determinar su incidencia en el uso de las secuencias idiomáticas en los aprendices de ELE; para valorar mejor el efecto de la estancia, hubiera sido más revelador comparar los resultados del grupo de suecos analizado con una estancia en España de unos cuatro meses con un grupo también de aprendices suecos de las mismas características que no hubiera realizado esta estancia; este contraste no se ha llevado a cabo.

Además, sería adecuado conocer las instrucciones que se les dieron a los estudiantes antes de las grabaciones de las negociaciones y grupos de discusión que se analizan: qué temas y cómo se les plantearon. Consideramos que la variable tema incide mucho en el tipo de unidades léxicas que emplean los participantes del corpus objeto de estudio, pero este factor no se tiene en consideración en el análisis.

Y concluimos esta reseña con una valoración de los resultados y su interpretación en relación con los conceptos de complejidad de una secuencia formulaica, su adaptación al destinatario y el grado de especialización que presenta. En la tesis doctoral se distinguen secuencias formulaicas más *complejas* que otras. A pesar de ello, no se detalla de forma sistemática en qué consiste esta complejidad: cuáles son los parámetros que se siguen para determinar que una unidad formulaica es más compleja que otras. Se apuntan algunos factores de forma algo dispersa: uso del subjuntivo, negaciones que abarcan varios sintagmas, alto contenido pragmático (cf. Moreno Teva 2012:186). Pero hablar de complejidad en relación con las secuencias formulaicas supone contradecir uno de los supuestos de los que parte la tesis que reseñamos: que las secuencias formulaicas se adquieren holísticamente. Por esta razón, no tiene sentido afirmar que es más compleja la secuencia *ir a INF* que *volver a INF*, o sea más que *es que*, por ejemplo. Si se postula que se adquiere toda la expresión formulaica en bloque, en principio el subjuntivo en el ejemplo del marcador *o sea* no debería ser un obstáculo, pues se procesaría holísticamente en su adquisición. La complejidad estructural que se aduce en el trabajo va en contra del supuesto de partida del procesamiento holístico de estas secuencias; esta contradicción manifiesta la falta de definición que hemos señalado al principio acerca del plano de análisis desde el que se estudian las secuencias lingüísticas objeto de este estudio.

En cuanto a la adaptación al destinatario, se comenta en el trabajo que los hablantes nativos utilizan secuencias formulaicas interregulativas para adaptar su discurso a los hablantes no nativos en las conversaciones exolingües; se necesitaría mostrar ejemplos de esta adaptación para saber en qué consiste, para entender cómo este tipo de secuencias interregulativas funcionan como estrategias de producción. Se afirma, además, que en los grupos de discusión exolingües los hablantes nativos se adaptan a un nivel menos complejo de uso de secuencias formulaicas, en el sentido de que son más breves y menos pragmáticas o

polifuncionales. Volvemos aquí a la discusión de la complejidad de estas unidades.

Por último, en relación con el grado de especialización de la unidad fraseológica, se caracterizan las secuencias formulaicas en el trabajo como pertenecientes al habla cotidiana, común o general en algunos casos, y como pertenecientes al lenguaje especializado en otros. Se afirma en la tesis, por ejemplo, que la combinación *en el mercado* pertenece al léxico específico de la negociación –y no de la actividad del grupo de discusión– (Moreno Teva 2012:108) o que la expresión *tener una base* no pertenece al léxico de la vida cotidiana (Moreno Teva 2012:111). En la bibliografía sobre el tema, se suele afirmar que las combinaciones estables de palabras son muy características del habla coloquial, forman parte del léxico coloquial, y también de discursos muy específicos. En este sentido, *en el mercado* sería un fenómeno más ligado a la terminología que a la fraseología: no es tanto la combinación de estas tres palabras lo que resulta específico en su uso sino el sentido que adopta el sustantivo *mercado* en un contexto empresarial. Falta analizar en el trabajo cuáles de las secuencias formulaicas identificadas forman parte del español para fines específicos y cuáles no.

En síntesis, a pesar de que en la tesis de doctorado reseñada se confirman las hipótesis de partida conviene que quienes la consulten tengan en cuenta las limitaciones que se recogen en el último apartado del trabajo: necesidad de realizar test previos a los hablantes para conocer más información de su persona (relación entre ellos, conocimiento de otros idiomas) y saber con exactitud su nivel de español antes de realizar su estancia en España. Como se ha dicho, se requeriría comparar la evolución en el uso de las secuencias formulaicas por parte de los hablantes no nativos durante su estancia en España con el uso de las secuencias formulaicas en hablantes no nativos que solo han sido objeto de enseñanza formal de español en Suecia. Además, se propone correlacionar la competencia formulaica de los hablantes no nativos con otros factores de la competencia lingüística o comunicativa como la complejidad sintáctica, la aptitud lingüística, la fluidez en la actuación o la motivación. Son todas ellas líneas futuras de investigación merecedoras de ser abordadas cuanto antes.

Carmen López Ferrero<sup>2</sup>

### Referencias citadas

- Corpas Pastor, Gloria (1997), *Manual de fraseología española*. Madrid: Gredos.  
Coseriu, Eugenio (1992), *Competencia lingüística: elementos de la teoría del hablar*. Madrid: Gredos.  
Erman, Britt & Warren, Beatrice (2000), “The idiom principle and the open choice principle”, *Text*, 20(1):29-62.  
Instituto Cervantes (2006), *Plan curricular del Instituto Cervantes*. Madrid: Instituto Cervantes y Ed. Biblioteca Nueva, S.L.

---

<sup>2</sup> Universitat Pompeu Fabra (Barcelona).

- Lewis, Michael (1993), *The Lexical Approach: The State of ELT and a Way Forward*. Hove: Language Teaching Publications.
- Wray, Alison (2002), *Formulaic Language and the Lexicon*. Oxford: Oxford University Press.
- Wray, Alison (2008), *Formulaic Language: Pushing the Boundaries*. Oxford: Oxford University Press.