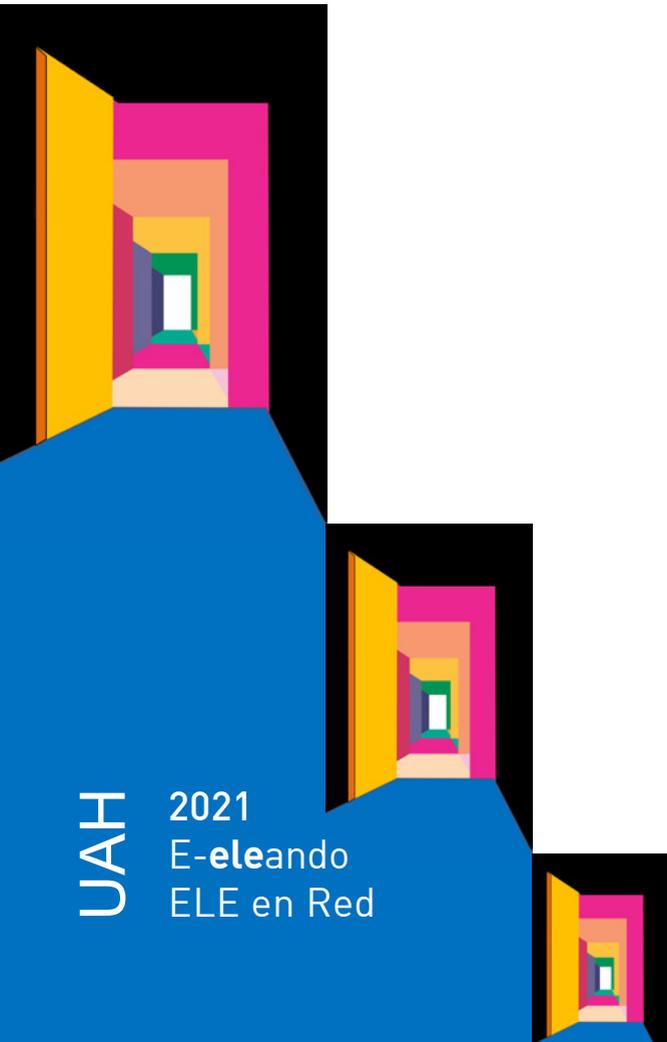


El lenguaje formulaico y su importancia en la adquisición/aprendizaje de segundas lenguas

Israel Escanilla Martínez



UAH

2021
E-eleando
ELE en Red

20

E-eleando

ELE en Red.

Serie de monografías y materiales para la enseñanza de ELE

Directora

Ana María CESTERO MANCERA (Universidad de Alcalá, España)

Secretaria

Arantxa GARCÍA DE SOLA RUBIO (Universidad de Alcalá, España)

Comité de Expertos

Marta ALBELDA MARCO (Universitat de València, España)

Fátima ÁLVAREZ LÓPEZ (Instituto Cervantes de Berlín
y Universidad Libre de Berlín, Alemania)

Elisa BORSARI (Universidad de Córdoba, España)

Laura CAMARGO FERNÁNDEZ (Universitat de les Illes Balears, España)

Sonia EUSEBIO HERMIRA (International House Madrid, España)

María del Carmen FERNÁNDEZ LÓPEZ (Universidad de Alcalá, España)

María Loreto FLORIÁN REYES (Escuela Oficial de Idiomas de Madrid, España)

Marta GARCÍA GARCÍA (Universität Göttingen, Alemania)

María JOSÉ GELABERT (Editorial EDINUMEN, España)

María Luisa GÓMEZ SACRISTÁN (Instituto Cervantes de Budapest, Hungría)

José Ramón GÓMEZ MOLINA (Universitat de València, España)

María Jesús MADRIGAL LÓPEZ (Alcalingua, Universidad de Alcalá, España)

Manuel MARTÍ SÁNCHEZ (Universidad de Alcalá, España)

María del Mar MARTÍN DE NICOLÁS MORENO (Universidad de Heidelberg, Alemania)

Beatriz MÉNDEZ GUERRERO (Universidad Autónoma de Madrid, España)

Elizabeth MIRAS PÁEZ (UDIMA, España)

Francisco MORENO FERNÁNDEZ (Universidad de Alcalá, España
y Universidad de Heildeberg, Alemania)

Zaida NÚÑEZ BAYO (Universidad de Alcalá, España)

Florentino PAREDES GARCÍA (Universidad de Alcalá, España)

Inmaculada PENADÉS MARTÍNEZ (Universidad de Alcalá, España)

Ana M.^a RUIZ MARTÍNEZ (Universidad de Alcalá, España)

Begoña SANZ SÁNCHEZ (Alcalingua, Universidad de Alcalá, España)

José SIMÓN GRANDA (Universidad de Alcalá, España)

Clara UREÑA TORMO (Universidad de Alcalá y Universidad Nebrija, España)

Reservados todos los derechos. Queda prohibida, salvo excepción prevista en la ley, cualquier forma de reproducción, distribución, comunicación pública y transformación de esta obra sin contar con la autorización de los titulares de propiedad intelectual. La infracción de los derechos mencionados puede ser constitutiva de delito contra la propiedad intelectual (arts. 270 y sigs., Código Penal). El Centro Español de Derechos Reprográficos (www.cedro.org) vela por el respeto de los citados derechos.

Esta publicación ha sido financiada por el Máster Universitario en Formación de Profesores de Español de la Universidad de Alcalá y por el Máster en Enseñanza de Español como Lengua Extranjera, estudio propio de la Universidad de Alcalá.



**MÁSTER UNIVERSITARIO EN FORMACIÓN DE
PROFESORES DE ESPAÑOL**



**MÁSTER EN ENSEÑANZA
DE ESPAÑOL COMO
LENGUA EXTRANJERA**

© Israel Escanilla Martínez

Editan: Máster Universitario en Formación de Profesores de Español de la Universidad de Alcalá y Máster en Enseñanza de Español como Lengua Extranjera, estudio propio de la Universidad de Alcalá
Colegio San José de Caracciolos
C/ Trinidad, 5 • 28801 Alcalá de Henares (Madrid, España)
Web: <http://www3.uah.es/meleuah/>

Universidad de Alcalá, Servicio de Publicaciones, 2021
Plaza de San Diego, s/n.º • 28801, Alcalá de Henares (España)
Web: uah.es

Difusión y redes: Arantxa García de Sola Rubio

El Máster Universitario en Formación de Profesores de Español de la Universidad de Alcalá y el Máster en Enseñanza de Español como Lengua Extranjera, estudio propio de la Universidad de Alcalá, no se hacen responsables de las ideas vertidas en este texto, así como tampoco lo son de la originalidad del mismo, ni de que en él se cumplan los cánones establecidos por la ley en reprografía.

e-ISSN: 2530-7606

<http://www3.uah.es/e-eleando>

El lenguaje formulaico y su importancia en la adquisición/aprendizaje de segundas lenguas

Israel Escanilla Martínez

Índice

1. INTRODUCCIÓN	3
2. EL ESTUDIO DE LO IDIOMÁTICO Y PREFABRICADO: FRASEOLOGÍA, FRASEODIDÁCTICA, GRAMÁTICA DE CONSTRUCCIONES Y LENGUAJE FORMULAICO.....	5
2.1. La fraseodidáctica	6
2.2. La gramática de construcciones	7
2.3. El lenguaje formulaico	10
2.4. Relación entre el lenguaje formulaico y la gramática de construcciones	11
3. EL CAMINO HACIA LO FORMULAICO Y LAS CONSTRUCCIONES	16
4. EL LENGUAJE FORMULAICO	20
4.1. Panorama histórico del análisis del lenguaje formulaico	22
4.2. Identificación del lenguaje formulaico	29
4.2.1. Métodos empíricos.....	30
4.2.2. Otros criterios.....	33
4.3. Taxonomía	35
5. FUNCIONES DE LAS FÓRMULAS: UN ENFOQUE ETNOLINGÜÍSTICO	37
6. ADQUISICIÓN DEL LENGUAJE FORMULAICO	43
6.1. Adquisición en niños (L1)	43
6.2. Aprendizaje en adultos (L2)	49
7. NUEVAS LÍNEAS DE INVESTIGACIÓN: LA LINGÜÍSTICA COMPUTACIONAL ..	56

8. CONCLUSIONES	58
REFERENCIAS BIBLIOGRÁFICAS	60
ANEXOS	63

1. Introducción

El propósito de la presente monografía es exponer las principales contribuciones realizadas desde el marco teórico del lenguaje formulaico —a partir de ahora, LF—. En primer lugar, se realizará una sucinta presentación sobre el estudio de lo idiomático y prefabricado, resaltando además del LF la fraseodidáctica, la gramática de construcciones y la relación entre esta última teoría y el LF. A continuación, se expondrá brevemente cómo bajo el paradigma del enfoque comunicativo de finales del siglo XX surge el gran interés por los fenómenos periféricos de la lengua y su enseñanza. Posteriormente, se explicarán los aspectos básicos del LF con el objetivo de comprender mejor este amplio enfoque. Así, se tratará su definición, la historia de su investigación, los métodos para su identificación y su taxonomía. La importancia de las fórmulas en la comunicación humana —ya sean estas abordadas desde la fraseología, la gramática de las construcciones o el LF— también será tratada, en este caso, desde un punto de vista etnolingüístico. El sexto capítulo contendrá un análisis de las principales aportaciones del estudio del LF en lo relativo a su adquisición en niños (L1) (L1) y en su aprendizaje en adultos (L2). Por último, se incluirá un apartado con las conclusiones.

Esta monografía se ha elaborado a partir de la consulta de los manuales, libros y artículos más relevantes en la investigación del LF. *Fundamentals of Formulaic Language* (2015) de David Wood, *Formulaic Language and the Lexicon* (2002) de Alison Wray y «Developments in the study of formulaic language since 1970: a personal view» (2007) de Andrew Pawley serán el principal material de apoyo en lo estrictamente referido al LF. Asimismo, se prestará especial atención a «Formulaic sequences: a drop in the ocean of constructions or something more significant?» (2016) de Andreas Buerki, artículo que busca trazar una relación sistemática entre el LF y la gramática de construcciones, y a «Regularity

and idiomaticity in grammatical constructions: the case of let alone» (1988), de Charles J. Fillmore *et al.*, artículo sustancial para el estudio de lo prefabricado, en concreto, para el asentamiento de la teoría de la gramática de construcciones. Además, «Aportaciones de la Etología Humana a los estudios lingüísticos: el caso de la Fraseología» (2015) de Olímpio de Oliveira Silva también será un material de apoyo principal. Si bien es cierto que el artículo de Olímpio de Oliveira Silva pone el foco en la aplicación de la etnolingüística a la fraseología, su estudio muestra el importante valor social y cultural de las fórmulas en la comunicación humana, ya sean estas particularmente investigadas desde la fraseología o desde un enfoque más amplio como el LF.

2. El estudio de lo idiomático y prefabricado: fraseología, fraseodidáctica, gramática de construcciones y lenguaje formulaico

El estudio de lo idiomático y de lo prefabricado ha sido y es abordado desde distintas corrientes y teorías. Podemos definir lo idiomático como toda combinación con un significado no interpretable a través del análisis de sus elementos. Es decir, lo idiomático como opuesto a lo composicional. Consecuencia de esta naturaleza, los modismos son idiosincrásicos, lo que suele manifestarse en una restricción geográfica. Por este motivo, los aprendientes de una lengua extranjera tienen numerosos problemas a la hora de comprender y adquirir este tipo de combinaciones. Enunciados como *refrescar la memoria* (colocación), *ser pan comido* (expresión), *al pie de la letra* (locución) o *cría cuervos y te sacarán los ojos* (paremia) son solo algunos ejemplos que muestran lo problemático que puede llegar a ser lo idiomático desde la perspectiva de un hablante extranjero. Esta dificultad no solo afecta a los aprendientes de una L2, sino también a los propios nativos. De acuerdo con Martí Sánchez, «muchos hablantes nativos actuales del español de España no tienen claro el sentido de [locuciones como] *Brindis al sol* ... o ignoran el refrán *El ojo, límpialo con el codo*» (2019: 5-6). A pesar de las dificultades que entraña su enseñanza tanto para el alumno como para el profesor, su comprensión y adquisición es necesaria para lograr una buena competencia comunicativa. Así lo recoge el *Plan Curricular del Instituto Cervantes* al incluir, en el inventario de tácticas y estrategias pragmáticas o en el de saberes y comportamientos socioculturales, el aprendizaje por distintos niveles de, por ejemplo, refranes, proverbios y locuciones o marcadores discursivos (conectores, reformuladores, estructuradores de la información, etc.), siendo

muchos de ellos unidades fraseológicas¹. En cuanto a lo prefabricado, podemos entenderlo como combinaciones de palabras con una distribución interna fija, preestablecida o ritualizada, pero sin la necesidad de ser idiomáticas, es decir, su comprensión también puede deducirse a través de la literalidad. En la obra pionera de la lingüística española *El discurs prefabricat* (2000), Vicent Salvador apunta lo siguiente sobre lo prefabricado:

El discurso repetido y prefabricado [subraya] la operatividad singular de unas unidades compuestas *ready-made* que constituyen una dimensión en el conjunto de la psicolingüística y sociolingüística. La idea de rutinización y de convencionalización —tanto en el nivel individual como en el social— planea sobre todo el área de los estudios fraseológicos actuales y es probablemente uno de sus alicientes más innegables ... El discurso prefabricado reúne un conjunto de trabajos que demuestran el creciente interés por esta parcela del conocimiento del lenguaje ... el carácter «prefabricado» ... se asigna a estas unidades que remiten generalmente a la idea de polilexicalidad necesaria de los compuestos de más de una palabra. (2000: 14)²

2.1. La fraseodidáctica

Una de las disciplinas más destacadas en la investigación de lo idiomático ha sido la fraseología. Altamente desarrollada a partir de la segunda mitad del siglo XX, la fraseología es la disciplina que estudia las expresiones fijas (colocaciones, locuciones, refranes...), siendo los frasemas o unidades fraseológicas sus principales unidades de análisis. En los años noventa emerge la fraseodidáctica, rama aplicada de la fraseología que demuestra un genuino interés por la enseñanza de estas unidades. Para comprender mejor esta rama de la fraseología, conviene puntualizar la distinción entre la fraseodidáctica y la didáctica de la fraseología. De acuerdo con González Rey, el origen de la fraseodidáctica se encuentra en la glotodidáctica o didáctica de lenguas, y pueden diferenciarse tres etapas principales en la historia de su constitución:

¹ https://cvc.cervantes.es/ensenanza/biblioteca_ele/plan_curricular/indice.htm

² [E]l discurs repetit i prefabricat [subratlla] l'operativitat singular d'unes unitats compostes *ready-made* que constitueixen una dimensió a l'ensens psicolingüística i sociolingüística. La idea de rutinització i de convencionalització —tant en el nivell individual com en el social— plana sobre tota província dels estudis fraseològic actuals i és probablement un dels seus al·licients més innegables... El discurs prefabricat aplega un conjunt de treballs que palesen l'interès creixent per aquesta parcel·la del coneixement del llenguatge ... el caràcter «prefabricat» ... s'assigna a aquestes unitats remet generalment a la idea de polilexicalitat necessària, do compostos de més d'un mot. (2000: 14)

[1] la *didáctica* de la fraseología, representando la didáctica una disciplina aplicada de la lingüística, dedicada a la enseñanza-aprendizaje de lenguas y siendo la fraseología el ámbito de las expresiones hechas de una lengua; [2] la didáctica de la *fraseología*, siendo la fraseología una disciplina lingüística y la didáctica un ámbito de aplicación; y [3] la fraseodidáctica, una rama de la fraseología aplicada como lo es también la fraseografía. (2012: 67-8)

Así, en la etapa de la *didáctica* de la fraseología aparecen las primeras manifestaciones de interés por la enseñanza de modismos, tanto en lenguas maternas como en lenguas extranjeras. En la etapa de la didáctica de la *fraseología* comienza el interés por parte de los especialistas (fraseólogos y paremiólogos) en términos relativos. Por último, la fraseodidáctica incumbe tanto a docentes como a especialistas, los cuales se preocupan por situar la didáctica de la fraseología en un lugar propio dentro de la fraseología (González Rey 2012: 68). González Rey afirma que hoy en día «la Fraseodidáctica se define no solo como la didáctica de la fraseología de una lengua sino también como la didáctica de toda una lengua a través de su fraseología» (2012: 67). La fraseodidáctica se centra así en la enseñanza de lo idiomático, como colocaciones, refranes, proverbios o dichos. Martí Sánchez señala que el objetivo principal de la fraseodidáctica no es otro que «proporcionar a los estudiantes reglas claras y seguras, acerca de una realidad que no puede generarlas» (2019: 8).

2.2. La gramática de construcciones

Surgida en los años ochenta, la gramática de construcciones es una de las principales teorías de la lingüística cognitiva, y ha sido durante la última década cuando su popularidad ha aumentado ostensiblemente. Tiene una fuerte base interdisciplinar, ya que recoge investigaciones de disciplinas como la psicolingüística, la neurolingüística o la lingüística computacional. La gramática de construcciones considera la lengua como un gran entramado de construcciones que se relacionan entre sí, poniendo especial énfasis en la importancia de la relación forma-significado. De esta manera, su principal unidad de análisis es la construcción. De acuerdo con Hernández Muñoz, «una construcción es un bloque sintáctico, semántico, pragmático, prosódico y cultural» (2019: 36). Sobre el concepto de construcción, Salvador escribe:

El concepto de *construcción* es fundamental dentro de la lingüística cognitiva y presenta una alternativa a la concepción de las reglas de combinación propia de mo-

delos gramaticales como el estructuralista o el generativista. Una construcción es un emparejamiento de forma y significado que no se ha de considerar ni completamente arbitrario ni completamente predecible ... El significado y el funcionamiento global de una construcción no se deriva directamente de la suma de los significados y el funcionamiento de sus componentes, sino que a menudo hay aspectos que se derivan de la interacción entre estos ... Las construcciones contienen información fonética, morfológica, léxica, sintáctica, semántica y también pragmática. (2000: 35)³

Por su parte, Fillmore define una construcción gramatical en los siguientes términos:

Por construcción gramatical entendemos cualquier patrón sintáctico al que se le asigne una o más funciones convencionales en un idioma, junto con lo que sea lingüísticamente convencionalizado en relación con su contribución para el significado o el uso de estructuras que lo contengan. (*apud* Salvador 2000: 35)⁴

Las primeras manifestaciones de la gramática de construcciones surgieron alrededor de los años ochenta en la lingüística inglesa de Estados Unidos como un intento de explicación sintáctica a los modismos, aportando un punto de vista distinto al de las teorías generativistas de tradición chomskiana (Calzado 2015: 3)⁵. Mientras la gramática generativa chomskiana se basa en el supuesto de que los seres humanos generamos las oraciones de una lengua aplicando una serie de reglas independientes de contexto, la gramática de construcciones concibe la gramática de una lengua como un sistema dinámico de unidades con una relación forma-función convencionalizada (Diessel 2013: 2).

³ El concepte de construcció és fonamental dins de la lingüística cognitiva i presenta una alternativa a la concepció de regles de combinació pròpia de models gramaticals com l'estructuralista o el generativista. Una construcció és emparellament de forma i significat, que no s'ha de considerar ni completament arbitrari ni completament predecible ... El significat i el funcionament global d'una construcció no es deriva directament de la suma dels significats i el funcionament del seus components, sinó que sovint hi ha aspectes que deriven de la interacció entre aquells ... Les construccions contenen informació fonètica, morfològica, lèxica, sintàctica, semàntica i també pragmàtica. (2000: 35)

⁴ By grammatical construction we mean any syntactic pattern which is assigned one or more conventional functions in a language, together with whatever is linguistically conventionalized about its contribution to the meaning or the use of structures containing it. (*apud* Salvador 2000: 35)

⁵ Como veremos más adelante, las teorías basadas en la gramática generativa de Chomsky han tenido problemas para dar respuestas al fenómeno de lo idiomático y de lo prefabricado en el lenguaje.

Durante esta primera etapa, marcó un punto de inflexión el artículo «Regularity and Idiomaticity in Grammatical Constructions: The Case of Let Alone» (1988), de C. J. Fillmore, P. Kay y M. C. O'Connor, el cual veremos más adelante. En una segunda etapa más reciente, se ha ampliado el concepto de construcción, convirtiéndolo en un «modelo uniforme de representación de todo el conocimiento gramatical» (Calzado 2015: 3). Los trabajos de los lingüistas W. Croft y A. Goldberg, y del psicólogo M. Tomasello están entre los más representativos en esta segunda etapa, mientras que el título de referencia más importante es *The Oxford Handbook of Construction Grammar* (2013), de Trousdale y Hoffmann (Hernández Muñoz 2019: 25). En este último título, Diessel apunta las dos premisas fundamentales de la teoría de la gramática de construcciones. En primer lugar, las estructuras sintácticas son unidades simbólicas, es decir, cada estructura tiene un significado. En segundo lugar, las construcciones se asocian entre sí a través de distintos tipos de vínculos, construyendo así un complejo entramado general de construcciones (2013: 2). Así, la lengua se entiende como un compendio o una red de construcciones de distinto tipo —construcciones sintácticas, construcciones léxicas...— que se combinan entre sí. Estas construcciones se configuran en la mente del hablante, siendo este es el último responsable en la creación del producto lingüístico.

En definitiva, la gramática de construcciones debe considerarse como un enfoque teórico basado en el uso que analiza las relaciones convencionales que existen entre forma y significado: «el principio básico de la gramática de construcciones ... es que las construcciones tradicionales —es decir, las correspondencias forma-significado— son las unidades básicas del lenguaje» (Goldberg 1995: 6)⁶. Su unidad de análisis es la construcción, y sus objetivos principales son comprender el proceso de adquisición de la lengua y su manejo. Según Calzado, aplicar un enfoque constructivista a la enseñanza de idiomas permite reunir gramática y comunicación, ya que bajo este modelo todas las construcciones gramaticales tienen su propia función comunicativa:

un enfoque construccionista basado en el uso, en el que toda la gramática es simbólica, puede ayudar a los alumnos a aprender (interpretar y producir) las construcciones del español dentro de un enfoque comunicativo de la enseñanza de la lengua. (2015: 12)

⁶ «the basic tenet of Construction Grammar ... is that traditional constructions —i.e., form-meaning correspondences— are the basics units of language» (Goldberg 1995: 6).

2.3. El lenguaje formulaico

Al igual que la gramática de construcciones, la teoría del LF surge de forma independiente en la lingüística anglosajona —*formulaic language*—. El LF hace referencia a toda aquella parte rígida y prefabricada del discurso (ya sea este oral u escrito), a aquellos conjuntos de palabras con una distribución fija o pres-tablecida que, en numerosas ocasiones, tienen un significado no deducible a partir del análisis de sus elementos. En este sentido, el LF, al igual que la fraseología, aborda el estudio de lo idiomático: «la fraseología es un término que se utiliza a menudo para referirse al estudio del lenguaje formulaico ... un término específico para un tipo particular de análisis del lenguaje formulaico» (Wood 2015: 2)⁷. Sin embargo, el LF no se limita únicamente al estudio de lo idiomático. Al igual que la gramática de construcciones, el LF aborda el estudio de lo ritualizado en el discurso desde una perspectiva más amplia, entrando dentro de lo conocido como *chunks*⁸. Dentro de esta categoría, no solo se encuentran las unidades prototípicas que estudia la fraseología, sino también todas aquellas estructuras cuyos elementos se encuentran altamente vinculados entre sí, pero sin la necesidad de tener un carácter idiomático. Así, el LF incurre en el estudio de fórmulas como *¿qué tal estás?, la verdad es que*, de marcadores discursivos cuyo carácter idiomático no resulta del todo claro como *en este caso o por un/otro lado*, o incluso de secuencias en las que se repite un mismo esquema como *me gusta el café café o es tonto tonto*. Debido a esta gran amplitud en el espectro de unidades analizadas, existen distintas definiciones sobre lo que es una secuencia formulaica y su identificación resulta compleja. Aun así, ambas cuestiones se abordarán en el cuarto capítulo de esta monografía. Además, aunque su objeto de investigación es claro —los conjuntos léxicos recurrentes en el discurso—, el LF no ha tenido apenas presencia en la teoría lingüística, especialmente durante el comentado periodo en el que dominaban las teorías estructuralistas y chomskianas:

Ubicar las SF [secuencias formulaicas] en la teoría lingüística ha sido tradicionalmente difícil, particularmente en la tradición estructuralista y chomskiana que domino el pensamiento durante el siglo XX. Esto se debe a que las SF se oponen a comprender la estructura lingüística como una combinación de palabras fuera

⁷ «phraseology is a term often used to refer to the study of formulaic language ... a specific term for a particular type of analysis of formulaic language» (Wood 2015: 2).

⁸ En inglés, un *chunk* es un conjunto de palabras almacenadas y procesadas como un todo. El psicólogo George A. Millner fue quien utilizó el término por primera vez en su artículo «The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information» (1956).

del lexicón de acuerdo a unas reglas combinatorias. Si uno intenta reconocer ... ejemplos de SF ... mientras mantiene una división entre reglas y palabras, pronto encontrará que las SF no encajan en ninguna categoría: las SF no son el producto de aplicar reglas generales a las palabras, ni se comportan de forma general como palabras únicas. (Buerki 2016: 15-6)⁹

Dos aspectos distintivos en la investigación del LF son la frecuencia y fluidez, esto es, la probabilidad de que un conjunto de palabras ocurra con asiduidad y la capacidad de los hablantes para producir secuencias sin vacilaciones ni interrupciones. El primer aspecto está asociado con los procesos de identificación del LF, mientras que el segundo se relaciona con el uso de secuencias formulaicas en aprendientes de una L2/LE. Estos dos aspectos se tratarán en más detalle en los capítulos tres y seis respectivamente. Por último, algunos de los mayores exponentes en la investigación del LF son David Wood y Alison Wray. Esta última es quien acuña el término *secuencia formulaica* —*formulaic sequence*— en la esencial *Formulaic Language and the Lexicon* (2002).

2.4. Relación entre el lenguaje formulaico y la gramática de construcciones

Tanto la gramática de construcciones como el LF comparten la idea de que la lengua es, en un alto grado, un proceso automatizado en el que la libertad de creación se encuentra más limitada de lo que cabría esperar. Ambas teorías adoptan un enfoque de la lengua basada en el uso, contraponiéndose así a las teorías derivadas de la gramática generativa, las cuales se apoyan más en la competencia gramatical innata. A pesar de compartir esta visión, se advierte complicado establecer unas fronteras lingüísticas completamente claras entre la gramática de construcciones y el LF. Perseguir este objetivo significaría adentrarse en un impreciso e inestable campo de investigación que, como veremos más adelante, ya resulta bastante difuso en términos taxonómicos. No ayuda tampoco la escasa bibliografía dedicada a este difícil propósito. Por el contrario,

⁹ The place of FSs [formulaic sequences] in linguistic theory has traditionally been a difficult one, particularly in regard to the structuralist and later Chomskyan traditions that dominated twentieth-century thinking. This is because FSs run counter to an understanding of linguistic structure as formed by combining words out of a store (the lexicon) according to general, combinatorial rules. If one tries to account ... examples of FSs ... while maintaining a division between rules and words, one soon finds that FSs do not fit in either category: they are not the products of general rules applying to words, and nor do they in general behave like single words. (Buerki 2016: 15-6)

esta carencia investigativa no existe en lo concerniente a la fraseología y la gramática de construcciones, cuya relación ha sido más investigada. Tal es el caso del estudio realizado por Dmitrij Dobrovol'skij en 2016 «Fraseología y Gramática de Construcciones», cuyo objetivo es aclarar que áreas de la fraseología pueden ser tratadas desde el marco teórico de la gramática de construcciones e, inversamente, que planteamientos de la gramática de construcciones pueden ser útiles para la fraseología. Wood afirma que el concepto de construcción tiene claramente un alto valor para todos aquellos interesados en establecer unas premisas teóricas para el estudio de lo formulaico:

Parece que los modelos con un enfoque basado en el uso y la gramática de la construcción cognitiva pueden ser útiles para elaborar una base teórica para el estudio del lenguaje formulaico... La gramática de construcciones... nos da un modelo lingüístico sobre la estructura del lenguaje que permite que el lenguaje formulaico ocupe fácilmente un lugar principal. (2015: 170)¹⁰

Uno de los pocos estudios que busca trazar una relación sistemática entre el LF y la gramática de construcciones es «Formulaic sequences: a drop in the ocean of constructions or something more significant?», realizado por Buerki en 2016. Buerki plantea en su artículo la difícil tarea de establecer si el LF es un fenómeno lo suficientemente autárquico como para no verse absorbido por las teorías constructivistas:

El autor considera si ... las secuencias formulaicas representan un fenómeno suficiente como para diferenciarlas y para justificarlas un estatus especial, o si por el contrario, es mejor estudiarlas como un conjunto más amplio de todas las construcciones que pueden encontrarse en la lengua. (2016: 15)¹¹

Para conseguir este objetivo, el autor realiza un estudio empírico: un extenso análisis de corpus, el cual contiene 29 millones de palabras procedentes de 63.075 entradas de la Wikipedia. Buerki establece que una secuencia de LF debe ser identificada como tal si esta aparece al menos dos veces por millón de

¹⁰ It appears that usage-based models of language acquisition and cognitive construction grammar may be of use in elaborating a theoretical basis for formulaic language study ... Construction grammar ... gives us a linguistic model of the structure of language which easily allows for formulaic language to take a primary place. (2015: 170)

¹¹ The author considers whether ... formulaic sequences represent a phenomenon that is sufficiently different to warrant special status or whether they might best be studied in terms of the larger set of all constructions found in language (2016: 15).

palabras (2016: 22). Dicha secuencia puede contener entre dos y nueve palabras (Buerki 2016: 22). Por último, y para obtener unos resultados más precisos, los conjuntos de palabras son catalogados como secuencias de LF en base a un criterio tipológico:

Mientras que las SF han sido categorizadas de muchas formas diferentes ... y las categorizaciones han sido criticadas por producir categorías discretas y no taxonomías completamente coherentes ... la tipología descrita a continuación resulta importante al ilustrar que las SF son más bien una colección heterogénea de elementos, sobre todo respecto a su grado de esquematización. Reconociendo que existen otras formas válidas de categorizar las SF, la tipología comprende lo siguiente: fórmulas ... colocaciones ... términos multi-palabra ... modismos ... proverbios ... y secuencias habituales. (2016: 24-5)¹²

Los resultados obtenidos llevan a Buerki a concluir que, en caso de querer categorizar las secuencias de LF dentro del marco teórico de la gramática de construcciones, estas deberían interpretarse como construcciones léxicamente sustantivas:

Si las SF se colocan dentro de un enfoque constructivista de la gramática, estas se encajan como construcciones sustantivas predominantemente léxicas, siendo su sustantividad la única característica formal que las distingue frente a otras construcciones. (2016: 29)¹³

Así, según Buerki, las secuencias de LF son una categoría que tiende a ser absorbida por la teoría de la gramática de construcciones:

Las secuencias formulaicas se disuelven en el océano de las construcciones ... Desde una perspectiva constructivista, es mejor considerar las SF como una etiqueta para una porción teóricamente definida de un gran fenómeno, más que un fenómeno por sí mismo – más diferente en grado que en tipo. Para los gramáticos cons-

¹² While FSs [formulaic sequences] have been categorised in very many different ways ... and categorisations have been criticised for producing neither discrete categories nor fully coherent taxonomies ... the typology outlined below is important in illustrating that FSs are a rather heterogeneous collection of items, not least with respect to their degree of schematicity. Recognising that there are other valid ways of categorising FSs, the typology comprises the following: formulas ... collocations ... multi-word terms ... idioms ... proverbs ... usual sequences. (2016: 24-5)

¹³ If FSs are placed within a constructionist approach to grammar, they slot in as predominantly lexically substantive constructions, their substantivity being the only formal distinguishing feature vis-à-vis other constructions. (2016: 29)

tructivistas, esto no es una sorpresa – su fluidez en los niveles de esquematización y su entremezclada omnipresencia en construcciones individuales es de lo que tratan las construcciones. Las SF simplemente forman parte de este universo. (2016: 29)¹⁴

Con todo, Buerki sostiene que el LF posee unas características peculiares respecto a la gramática de construcciones. El procesamiento del lenguaje, la adquisición de una L1, y el aprendizaje y la comunicación en una L2 son rasgos psicolingüísticos propios en la investigación del LF, y que lo hacen diferenciarse de las teorías constructivistas:

(el procesamiento del lenguaje, la adquisición de la L1 y la L2, y la comunicación) indican que las SF tienen un significado teórico que las distingue de otras construcciones. Esto, a su vez, conlleva a plantearse unas cuestiones diferentes sobre la investigación de las SF en comparación con las construcciones, y las otorga un significado especial a pesar de su delineación algo artificial como fenómeno. Desde este punto de vista, las SF son diferentes y plantean unas preguntas especiales. Estas preguntas resultan importantes para una comprensión completa sobre el funcionamiento del lenguaje. (2016: 29)¹⁵

En conclusión, el estudio de Buerki plantea la posibilidad de integrar las secuencias de LF dentro del marco teórico de la gramática de construcciones. En palabras del autor, esta integración «no solo es posible sino deseable en muchos aspectos» (2016: 30)¹⁶. No obstante, Buerki afirma que nuestra visión sobre el LF y la gramática de construcciones no debe limitarse únicamente a establecimiento de fronteras formales entre ambos fenómenos. Los aspectos psicolingüísticos mencionadas anteriormente también deben tenerse en cuenta:

¹⁴ So formulaic sequences dissolve into the ocean of all constructions ... From a constructionist perspective at least, FSs are therefore best seen as a convenient label for a theoretically defined portion of a larger phenomenon rather than a self-contained phenomenon – different in degree, rather than type. To construction grammarians, this is not a surprise – the fluidity of levels of schematicity and their ubiquitous intermixing in single constructions is what constructions are all about and FSs, where acknowledged, were simply taken to be part of this universe (2016: 29).

¹⁵ (language processing, first language acquisition, second language acquisition and communication) indicate that FSs bear theoretical significance that marks them out as different from other constructions. This in turn results in particular and distinct research questions being asked of FSs compared to other constructions and imbues them with special significance notwithstanding their somewhat artificial delineation as a phenomenon. From this point of view, FSs are different and need special questions asked of them. Furthermore, as outlined, these questions are important for a full understanding of the workings of language. (2016: 29)

¹⁶ «is not only possible but desirable in many respects» (2016: 30).

Quedan otras formas de ver las SF y las construcciones ... Mientras que hoy en día pocos defienden adoptar una postura basada en la forma como característica principal o exclusiva de la estructura lingüística, sería fascinante investigar cómo se relacionan las SF y las construcciones en términos de representación psicolingüística y de procesamiento ... Dadas las conexiones y la falta de claridad entre las SF y otras construcciones que se muestran en este estudio ... resultaría sorprendente sin embargo que el resultado general de una visión psicolingüística sobre las SF y sobre otras construcciones concluyera que son fenómenos completamente separados. (2016: 30)¹⁷

¹⁷there remain other ways of looking at FSs and constructions ... While taking form as the primary or exclusive characteristic of linguistic structure is today advocated by few, it would be fascinating to investigate how FSs and constructions relate in terms of psycholinguistic representation and processing ... Given the connections and the blurring between FSs and other constructions shown in this study ... it would however seem surprising if the overall outcome of a psycholinguistically focused look at FSs and other constructions were to conclude that they are entirely separate phenomena. (2016: 30)

3. El camino hacia lo formulaico y las construcciones

Durante las tres últimas décadas, el peso de lo idiomático y prefabricado ha cobrado una gran importancia en la enseñanza de lenguas extranjeras. En gran parte, esto se debe al cambio en el paradigma lingüístico y educativo en la segunda mitad siglo XX. El auge del enfoque comunicativo como modelo didáctico durante los ochenta y noventa puso el foco de la enseñanza en el aprendiente, pretendiendo capacitar a este para hacer un uso real de la lengua. Los modelos educativos previos tenían una fuerte base estructuralista, y se fundamentaban en la enseñanza de la gramática, ya que consideraban que la competencia lingüística era el aspecto básico para lograr un alto grado de competencia en una lengua extranjera.

La competencia lingüística —o competencia gramatical— es la capacidad que posee un hablante para producir y comprender enunciados gramaticales, es decir, enunciados que se adapten a las reglas gramaticales de la lengua¹⁸. Estos enfoques estructuralistas comenzaron de la mano de Saussure con su obra *Curso de lingüística general* (1916). Posteriormente, el trabajo de Saussure fue continuado, entre otros muchos, por lingüistas europeos como André Martinet o Emilio Alarcos, o americanos, como Bloomfield o Hockett.

Por el contrario, y sin desestimar la importancia de la competencia lingüística, el enfoque comunicativo busca una competencia real del hablante, es decir, busca capacitar a los aprendientes para hacer un uso instrumental de la lengua que sirva como vehículo de interacción con sus semejantes. De esta manera, la lengua

¹⁸ El concepto competencia gramatical fue propuesto por N. Chomsky en *Estructuras sintácticas* (1957), y resulta clave en la tradición de la gramática generativa (*Diccionario de términos clave de ELE*).

ya no era considerada únicamente como un conjunto de elementos, estructuras y reglas, por lo que el enfoque comunicativo fue considerado un avance frente a los modelos estructuralistas previos iniciados por Saussure a comienzos del siglo XX. Si bien es cierto que el enfoque comunicativo ha ido evolucionando y dando paso a otros modelos, como el enfoque por tareas, es necesario destacar que fue bajo este paradigma educativo desde donde se comenzó a prestar realmente atención al fenómeno de lo idiomático y prefabricado en la enseñanza de lenguas extranjeras.

En este sentido, el enfoque léxico propuesto por Michael Lewis en *The Lexical Approach* (1993) es un claro ejemplo de la evolución que produjo el enfoque comunicativo en las metodologías para la enseñanza de lenguas extranjeras. Lewis expone la necesidad de abandonar los modelos educativos estructuralistas y aplicar un enfoque de la lengua basado en el uso, el cual nos obliga a replantearnos el papel del vocabulario y la gramática en la enseñanza de una LE:

Esta palabra [estructura], que una vez fue fundamental para toda la enseñanza de idiomas, ahora está un tanto pasada de moda. Se utiliza con mayor frecuencia para hablar sobre *syllabus* estructuralistas. Hasta hace 20 años, estos *syllabus* no se cuestionaban. La enseñanza de idiomas se basaba en la suposición de que los estudiantes primero necesitaban dominar la estructura de unas determinadas oraciones. Se suponía que una vez que las dominaban, posteriormente podían aprender a «rellenar los huecos» en oraciones estructuralmente correctas, ampliando su vocabulario y simplemente insertando nuevas palabras en los espacios de las oraciones. Las estructuras que fueron, y que en gran medida todavía son, percibidas como esenciales para la enseñanza de idiomas fueron únicamente las de oraciones bien formadas en inglés. (1993: 3)¹⁹

El enfoque léxico de Lewis implica un papel menos preponderante para la gramática en el aula, —al menos, hasta niveles intermedio-altos— y otorga al léxico un rol principal: «la lengua consiste en un léxico gramaticalizado, no en una

¹⁹This word [structure], once central to all language teaching, is now somewhat out of fashion. It is most often used to talk about structural syllabuses. Until 20 years ago such syllabuses were unchallenged. Language teaching was based on the assumption that students first needed to master particular sentence frames — structures. The assumption was that once they had mastered these, they could subsequently learn how to ‘fill the gaps’ in structurally correct sentences, by expanding their vocabulary and simply inserting new words into slots within the sentences. The structures which were, and to a large extent still are, perceived as central to language teaching were precisely, and exclusively those of well-formed English sentences. (1993: 3)

gramática lexicalizada» (1993: 51)²⁰. De acuerdo con Lewis, conocer una lengua no consiste únicamente en saber cómo formar oraciones gramaticales, sino en cómo transformar esas oraciones en actos comunicativos (1993: 11). Este planteamiento fue ignorado en las metodologías basadas en la teoría chomskiana:

Aprender un idioma era sinónimo de dominar las estructuras del mismo, es decir, lograr la competencia. La suposición tácita era que una vez que se dominabas el sistema, podías generar la frase que quisieses. Que los hablantes nativos poseían una «competencia de hablante nativo» era un axioma. El pensamiento de Chomsky era completamente teórico y su abstracto modelo tuvo una influencia considerable. Hoy en día, para la mayoría de los lingüistas y metodólogos, es un modelo ampliamente desacreditado. (Lewis 1993: 4)²¹

El enfoque léxico de Lewis ha sido importante para demostrar, dentro del paradigma del enfoque comunicativo, la importancia que tiene en la enseñanza de una LE los *chunks*, concepto recurrente la investigación del LF. Estos *chunks* son cadenas prefabricadas de palabras a las que recurrimos de forma casi automática para realizar actividades comunicativas.

Muchos respaldan (...) [el] punto de vista de que los hablantes nativos retienen (de forma redundante) muchos elementos prefabricadas y formulaicos (*chunks*). La fluidez se logra en gran medida mediante la combinación de *chunks*, lo que reduce la dificultad de procesamiento. Esto se relacionan con la forma en que almacenamos y recuperamos el lenguaje de forma natural, pero claramente existen enormes implicaciones para el aprendizaje de idiomas... Es esencial aumentar la conciencia de los estudiantes sobre los fragmentos de los que se compone un discurso coherente o un texto escrito. (Lewis 1993: 121)²²

²⁰ «Language consists of grammaticalised lexis, not lexicalised grammar» (1993: 51).

²¹ Learning a language was synonymous with mastering the structures of the language, that is, achieving competence. The tacit assumption was that once you had mastered the system, you could generate any sentence you wished. It was axiomatic that native speakers possessed 'native speaker competence'. Chomsky's thinking was entirely theoretical and his abstract model exercised considerable influence. Nowadays, for most linguists and methodologists, it is largely discredited. (Lewis 1993: 4)

²² Many commentators endorse (...) [the] view that native speakers retain (redundantly) many pre-fabricated formulaic items — chunks. Fluency is achieved largely by combining chunks, reducing processing difficulty. These matters relate to the way we naturally store and retrieve language but clearly there are enormous implications for language learning ... it is essential to heighten students' awareness of the bits of which coherent discourse or written text is composed. (Lewis 1993: 121)

Por último, junto con el auge del enfoque comunicativo debe resaltarse que el camino hacia lo formulaico y las construcciones prosperó gracias al desarrollo en paralelo de distintas disciplinas y campos de investigación:

Esta nueva concepción de la enseñanza / aprendizaje de la LE [el enfoque comunicativo] recoge las aportaciones de varios campos de investigación, tales como la lingüística funcional británica (por ejemplo, J. Firth y M. A. K. Halliday), la sociolingüística estadounidense (por ejemplo, D. Hymes, J. Gumperz y W. Labov) y la filosofía del lenguaje o pragmalingüística (por ejemplo, J. Austin y J. Searle). (*Diccionario de términos clave de ELE*)

Los avances en estas y otras disciplinas han sido recogidos tanto por la gramática de construcciones como por el LF. Desde la gramática de construcciones, se otorgó gran importancia al estudio de novedosos aspectos como la semántica, la pragmática o la cognición, los cuales no habían sido suficientemente considerados en las teorías lingüísticas previas sobre la idiomática del lenguaje. A este respecto, Goldberg afirma: «... La gramática de construcciones ha crecido en gran medida desde el trabajo realizado en la semántica ... y desde un enfoque experiencial de la lengua» (1995: 6-7)²³. Por otra parte, el LF se encuentra altamente influenciado por disciplinas como la psicolingüística, la sociolingüística, la filosofía del lenguaje o la pragmalingüística. De acuerdo con Wray,

un modelo lingüístico que pretenda explicar el lenguaje formulaico no solo ha de tener en cuenta la estructura lingüística, sino también la motivación de su uso, es decir, cómo el comportamiento social y psicológico de los hablantes determina su comportamiento lingüístico. (*apud Contreras Izquierdo y Rey Arranz 2018: 23*)

²³ «... Construction Grammar has grown largely out of work on frame semantics ... and an experientially based approach to language» (1995: 6-7).

4. El lenguaje formulaico

En primer lugar, es necesario remarcar que no existe una definición comúnmente seguida de LF. Aun así, la idea subyacente es clara: existen secuencias de palabras —*multi-word sequences*— que tienden a aparecer juntas en los textos, escritos u orales. El hablante, nativo o no nativo, almacena estas unidades de forma holística, es decir, como un conjunto. Como veremos en el tercer capítulo, el LF tiene un papel fundamental tanto en la adquisición de la L1 en niños, como en el aprendizaje de la L2 en adultos.

En *Formulaic Language and the Lexicon* (2002) Wray define una secuencia de LF de la siguiente manera:

una secuencia, continua o discontinua, de palabras u otros elementos que es, o parece estar, prefabricada, es decir, almacenada y recuperada entera de la memoria en el momento de su uso, en vez de ser generada o analizada por la gramática del lenguaje (2002: 9)²⁴.

Por el mismo camino, David Wood sostiene lo siguiente en cuanto a la definición de LF:

El consenso general sobre la definición de lenguaje formulaico parece afirmar que los elementos que la conforman son: 1 Palabras múltiples [;] 2 Con un solo significado funcional [.] 3 Prefabricados o almacenados y recuperados mentalmente como una sola palabra (2015: 3)²⁵.

²⁴ a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar. (2002: 9).

²⁵ The general consensus on a definition of formulaic language seems to be that the items will be:

Como Wray expone, las secuencias de FL pueden ser discontinuas. Esto implica que una secuencia formulaica puede tener un «hueco» —*slot*— que acepte la inclusión de distintas palabras —*fillers*—. Por ejemplo, en la secuencia *hace cincuenta años que*, el *slot* sería, claramente, la palabra *cincuenta*, ya que la estructura recurrente es *hace X años que*, mientras que *cincuenta* es la palabra concreta —*filler*— que el hablante necesita utilizar en ese momento transmitir una determinada información. Otros ejemplos, son la secuencia *no solo X, sino Y*, o la secuencia comparativa inglesa *the X, the better*. De acuerdo con Wray, estas son las secuencias semifijadas —*semi-fixed sequences*—: «las secuencias semifijadas ... contienen ranuras o huecos para una variedad de material obligatorio y opcional para ser insertado» (2002: 34)²⁶. Asimismo, Wray afirma que estos huecos ofrecen la posibilidad al hablante de producir un discurso algo menos restringido, es decir, proporcionan un cierto grado de libertad para que el discurso no resulte tan mecánico: «El uso de secuencias formulaicas siempre impone restricciones a la novedad, y esto generalmente se compensa insertando los elementos apropiados de clase abierta en los huecos» (2002: 222)²⁷. A partir de la definición de Wray, podemos asumir la existencia de secuencias con un grado de complejidad difícilmente explicables desde las teorías de la gramática generativa o transformacional. De acuerdo con Wray, estas secuencias de LF se almacenan, es decir, no pasan por un filtro de análisis antes de ser adquiridas. Posteriormente, son recuperadas de la memoria y reproducidas como un completo. Se deduce así, que la gramática generativa no es la única teoría plausible para explicar los productos lingüísticos de los hablantes, y que las secuencias de LF demuestran la existencia de otros procesos cognitivos a la hora de utilizar la lengua. A este respecto, Moreno Teva afirma en su tesis, *Las secuencias formulaicas en la adquisición de español L2* (2012), lo siguiente:

a partir de teorías gramaticales basadas más en la actuación que en la competencia lingüística, se ha ido encontrando cada vez más evidencia de la existencia de unidades léxicas mayores y con cierta complejidad interna. Estas unidades también se pueden combinar con otras unidades grandes o pequeñas, reduciendo así la actividad gramatical necesaria para expresar un mensaje completo. La existencia de

1 Multi-word [;] 2 Have a single meaning of function [;] 3 Be prefabricated or stored and retrieved mentally if a single word. (2015: 3)

²⁶ «semi-fixed sequences ... contain slots for a variety of compulsory and optional material to be inserted» (2002: 34).

²⁷ «Using formulaic sequences always imposes restrictions on novelty, and this is usually compensated for by inserting the appropriate open class items into the slots» (2002: 222).

unidades lingüísticas mínimas no queda por ello excluida, pero se otorga una la flexibilidad mayor a la expresión y a la comprensión del lenguaje. (2012: 18)

Por último, y para terminar con los aspectos conceptuales, es necesario apuntar la diferencia que existe entre LF y secuencia formulaica, la cual puede dar lugar a equívoco. Una secuencia formulaica, como hemos visto, se refiere a un conjunto de palabras que son tratadas como una unidad, mientras que LF es un término incontable que alude a esas secuencias valorándolas como un colectivo que representa en su totalidad un lenguaje prefabricado. En palabras de Wood, «las secuencias formulaicas generalmente son usadas para referirse a uno de estos elementos, mientras que el lenguaje formulaico es el sustantivo incontable que hace referencia a estos elementos como un colectivo» (2015: 2)²⁸.

4.1. Panorama histórico del análisis del lenguaje formulaico

Las primeras aportaciones relevantes para el estudio del LF aparecieron a principios del siglo XX. Esta primera etapa está caracterizada por la amplia variedad de disciplinas que contribuyeron al estudio de lo formulaico. La gran recapitulación realizada por Andrew Pawley en «Developments in the study of formulaic language since 1970: a personal view» (2007) nos sirve para trazar un panorama histórico sobre el estudio del LF. De acuerdo con el lingüista australiano, en los años setenta ya podían distinguirse, al menos, nueve disciplinas diferentes que abordaban algún aspecto del LF:

En 1970, académicos de al menos nueve disciplinas diferentes —estudios literarios, estudios sobre el folclore, antropología social, neurología, psicología experimental, psicología de la educación, microsociología, enseñanza del inglés como lengua extranjera y lexicografía— habían realizado importantes investigaciones sobre aspectos del lenguaje formulaico. (2007: 3)²⁹

²⁸ «formulaic sequences generally used to refer to one such item, formulaic language is the uncountable noun referring to these items as a collective» (2015: 2).

²⁹ By 1970, it is far to say, scholars in at least nine different disciplines — literary studies, folklore studies, social anthropology, neurology, experimental psychology, educational psychology, micro-sociology, the teaching of English as a foreign language and lexicography — had done significant research on aspects of formulaic language. (2007: 3)

Sin embargo, las investigaciones previas a los años setenta no prestaron la suficiente atención al LF. El predominio de las teorías basadas en la sintaxis fue una de las causas. Pawley comenta al respecto:

En la década de 1960, las expresiones formulaicas se consideraban generalmente marginales en la lingüística descriptiva o teórica. Varias mentalidades sostuvieron esa opinión. Primero, se consideró que el discurso formulaico constituía solo una pequeña parte de la producción normal de los hablantes nativos. Así, todo el énfasis recaía en el poder de la sintaxis para generar un número infinito de oraciones. Aprendimos a cantar el mantra de que «la mayoría de las oraciones que pronuncian los hablantes son innovadoras u originales». (2007: 11)³⁰

... Las fórmulas bien formadas (como proverbios, clichés y muchas colocaciones) no interesan a los gramáticos porque son simplemente expresiones convencionales o comunes. De lo que se trata la gramática es de las reglas y la estructura, y no se preocupa por las nociones de «uso establecido» o «frecuencia de uso» ... (12)³¹

Asimismo, Wood indica otras dos causas importantes para esta indiferencia. En primer lugar, las limitaciones tecnológicas propias de la época para realizar amplios análisis de corpus. En segundo lugar, la carencia de conocimientos lingüísticos necesarios en algunas de las disciplinas (2015: 4). A pesar de estos inconvenientes, algunas de las disciplinas mencionadas anteriormente aportaron unas primeras investigaciones muy interesantes para advertir la importancia de las construcciones en la lengua.

Dentro de los antropólogos y folcloristas, sobresalen los estudios acerca de los rituales orales. Pawley destaca *Coral Gardens and Their Magic* (1935), del antropólogo Bronisław Malinowski. En esta investigación etnográfica sobre los habitantes de las islas Trobiand de Nueva Guinea, Malinowski observa como la lengua oral de los rituales mágicos está compuesta por un LF, con un tono y ritmo característico, y con unos significados no literales. Pawley resalta como en la

³⁰ In the 1960s formulaic expressions were generally regarded as marginal to the proper subject matter of descriptive or theoretical linguistics. Several mindsets sustained that view. First, formulaic speech was considered to constitute only a small part of the normal output of native speakers. All the emphasis then was on the power of syntax to generate an infinite number of sentences. We learnt to chant the mantra that “most of the sentences that speakers utter are novel”. (2007: 11)

³¹ ... well- formed formulae (such as proverbs, clichés and many restricted collocations) are of no interest to grammarians because they are merely conventional or commonplace expressions. Grammar is about rules and structure, and gives no place to the notions of “established usage” or “frequency of use” ... (12)

investigación de Malinowski pueden observarse características prototípicas del LF: «Observe, una vez más, una combinación de características atribuidas a las fórmulas: palabras fijas, entonación distintiva, ritmo distintivo y función pragmática separada del significado literal» (2007: 6)³². Asimismo, en 1962 Dell Hymes —quien pasa por ser el que acuñó el término *competencia comunicativa*— escribe «The Ethnography of Speaking», un ensayo de referencia antropológica sobre la etnografía del habla y los rituales de encuentro. En él, Hymes afirma que

una gran proporción de la conducta verbal consiste en patrones recurrentes, ... [incluyendo] la gama completa de expresiones que adquieren un significado convencional para un individuo, un grupo o toda una cultura. (*apud* Pawley 2007: 7)³³

La filosofía y la sociología aportaron interesantes investigaciones sobre el uso del lenguaje cotidiano como una estrategia en la interacción. En los años sesenta el sociólogo Erving Goffman analizó cómo la presencia de estructuras preestablecidas en el discurso —construcciones— estaban determinadas por las normas sociales que regían el comportamiento de los individuos en público (Pawley 2007: 7). Desde la filosofía del lenguaje, las aportaciones más relevantes son las realizadas por John Austin y John Searle, también en los sesenta. Ambos autores desarrollaron la teoría de los «actos del habla», la cual sostiene que todo enunciado tiene una función discursiva, es decir, todo enunciado, además de presentar una información, representa una acción o intención del emisor (saludar, disculparse, criticar, acusar...). De acuerdo con Pawley, lo interesante de esta teoría es que estos enunciados, normalmente, adoptan la forma de expresiones convencionales o prefijadas, es decir, son LF (2007: 7).

En la neurología y la neuropsicología, las investigaciones más aprovechables para el estudio del LF son aquellas que examinan la localización de las funciones del lenguaje en el cerebro. El estudio más trascendente hasta la fecha es el iniciado por Paul Broca en 1860. Tras el análisis de pacientes con daños en el cerebro, Broca demostró el papel fundamental del hemisferio izquierdo en el procesamiento y producción del lenguaje. Pawley menciona cómo posteriormente el neurólogo inglés John Hughlings Jackson descubre que ciertos tipos severos de afasia en el área Broca dejan a los individuos sin la posibilidad de producir

³² «Notice, once again, a combination of features attributed to formulae: fixed words, distinctive intonation, distinctive rhythm, and pragmatic function separate from literal meaning» (2007: 6).

³³ «a vast proportion of verbal behaviour consists of recurrent patterns, ... [including] the full range of utterances that acquire conventional significance for and individual, group or whole culture» (*apud* Pawley 2007: 7).

un lenguaje proposicional, pero, sin embargo, siguen teniendo la capacidad de producir un lenguaje automatizado (2007: 8). Dicho de otra forma, las personas gravemente afectadas por una afasia de Broca pierden la capacidad de producir un lenguaje significativo que contenga palabras e informaciones con sentido. Por el contrario, mantienen la capacidad de producir un lenguaje mecánico basado en la memoria. Por ejemplo, las personas afectadas por una grave afasia de Broca podrían contar los días de la semana, recitar un poema, cantar una canción... es decir, podrían producir, en esencia, un LF.

Por su parte, los gramáticos de principios y mediados de siglo XX también reconocieron la importancia de las expresiones convencionales en el lenguaje. Otto Jespersen, lingüista danés especializado en gramática y en la enseñanza del inglés como lengua extranjera, realiza en su obra más conocida, *Language: Its Nature, Development and Origin* (1922), la distinción entre expresiones «libres» y expresiones «fijas» (Pawley 2007: 9). Sin embargo, no fue hasta finales de los años sesenta cuando los gramáticos comenzaron realmente a considerar las dificultades que provocaban las construcciones para los modelos de lenguaje basados en un límite claro entre la sintaxis y el léxico. De acuerdo con Pawley, a partir de estos años aparecieron trabajos más relevantes sobre la gramática y lo idiomático, siendo los de Makkai, Chafe y Weinreich los más destacables (2007: 9).

En *Idiom Structure in English* (1972), Adam Makkai presenta la teoría de estratificación del lenguaje y realiza una clasificación de los modismos —*idioms*—, dividiéndolos en *lexemic idioms* y *sememic idioms* (1972: 117). Sin embargo, el meticuloso trabajo de Makkai no estuvo exento de críticas en varios de sus postulados. Hudson comentó respecto a la taxonomía propuesta por Makkai:

La distinción entre *lexemic* y *sememic idioms* es más difícil de justificar, porque está claro que en AMBOS casos la cadena de morfemas tiene que ser tratada como una única *sememic unit*, y no sería suficiente para un «*lexemic*» *idiom* que este se unificara solo en el estrato «*lexemic*». De hecho, parece ser casi una cuestión de definición que las unidades de sintaxis (lexemas y grupos de lexemas) DEBEN corresponder a unidades de la semántica (*sememes* y grupos de *sememes*). (1975: 77)³⁴

³⁴ The distinction between lexemic and semantic idioms is harder to justify, because it is clear that in BOTH cases the string of morphemes has to be treated as a single sememic unit, and it would not be enough for a 'lexemic' idiom to be unified only on the 'lexemic' stratum. Indeed, it seems to be almost a matter of definition that the units of syntax (lexemes and groups of lexemes) MUST correspond to units of semantics (sememes and groups of sememes). (1975: 77)

Makkai aseguraba que los modismos debían tener un significado, además de idiomático, literal. Esta posición también fue puesta en entredicho:

Y, en tercer lugar, ¿por qué DEBE ser relevante la posibilidad de una interpretación literal para la definición de modismos? Después de todo, presumiblemente el interés principal de los modismos (como sea que los definamos) está en las relaciones entre sus estructuras internas y su significado, y estas relaciones son las mismas, sea posible o no una interpretación literal. (Hudson 1975: 75)³⁵

Con «Idiomaticity as an Anomaly in the Chomskyan Paradigm» (1968), Wallace Chafe, lingüista estadounidense conocido por sus críticas a la teoría generativista chomskiana, remarca la importancia de la semántica como componente fundamental del lenguaje. En su artículo, Chafe pone en relieve las dificultades que surgen al intentar explicar lo idiomático, como los modismos, desde la teoría chomskiana. Un cambio hacia un modelo basado más en la semántica que en la sintaxis es necesario para poder dar respuesta al fenómeno de lo idiomático:

Un cambio de un punto de vista sintáctico generativo a uno semántico generativo no es un paso pequeño, y requerirá la solución de muchos problemas que han sido vagamente o nada previstos en este momento. Los modismos, creo, proporcionan un tipo de evidencia, aunque no la única, de que se debe dar ese paso si la lingüística ha de aumentar significativamente el dominio de los fenómenos que puede explicar. La ingeniosa resolución de problemas dentro del paradigma chomskiano que ahora domina la literatura lingüística solo puede llegar hasta cierto punto, y ahora se encuentra con dificultades como las que se han descrito en este artículo. (Chafe 1968: 126-7)³⁶

Por último, en «Problems with the analysis of idioms» (1969), Uriel Weinreich se preocupa por identificar y diferenciar lo idiomático en la estructura de la lengua:

³⁵ And thirdly, why SHOULD the possibility of a literal interpretation be relevant to the definition of idioms? After all, presumably the main interest of idioms (however we define them) is in the relations between their internal structures and their meaning, and these relations are the same whether a literal interpretation is possible or not. (Hudson 1975: 75)

³⁶ A shift from a generative syntactic to a generative semantic point of view is no small step, and it will necessitate the solution of many problems only dimly or not at all foreseen at the moment. Idioms, I think, provide one kind of evidence, although not the only kind, that such a step must be taken if linguistics is to increase significantly the domain of phenomena which it can explain. The ingenious problem-solving within the Chomskyan paradigm which now dominates the linguistic literature can go only so far, and is now running into difficulties such as those which have been described in this paper. (Chafe 1968: 126-7)

El primer problema que tenemos ante nosotros es identificar la idiomática en la estructura del lenguaje, ver qué son las expresiones idiomáticas y en qué se diferencian de las expresiones no idiomáticas; el segundo es ver si las expresiones idiomáticas se pueden «disolver», mediante procedimientos analíticos apropiados, en constituyentes no idiomáticos, y qué obstáculos se interponen en el camino de tal disolución. (Weinreich 1969: 26)³⁷

Weinreich sostiene que un modismo debe ser «una expresión compleja cuyo significado no puede derivarse del significado de sus elementos» (1969: 26)³⁸ y, al igual que Chafe, resalta los problemas que conlleva abordar el estudio de lo idiomático desde las teorías de origen chomskiano:

Finalmente, para una lingüística que se preocupe intensamente por la productividad, en el enfoque chomskiano la idiomática representa un escollo teórico básico; porque bajo la idea de idiomática nos ocupamos de estructuras complejas que pueden ser reconocidas y analizadas, pero no generadas naturalmente por ninguna maquinaria explícita propuesta hasta este momento. (1969: 22)³⁹

Como puede observarse, en los años sesenta los gramáticos ya comenzaron a abordar más detenidamente el fenómeno de las construcciones, centrándose un buen número de ellos en el análisis de los modismos, un tipo de unidad dentro del LF. A pesar de que una parte de estas investigaciones han sido refutadas a lo largo del tiempo —véase el caso de *Idiom Structure in English* de Adam Makkai—, no dejan de tener un valor lingüístico e histórico para el estudio del LF. El punto de inflexión ocurre en los años setenta, momento a partir del cual una mayor cantidad de lingüistas y lexicógrafos se adentraron de lleno en el estudio de las construcciones.

Destaca en esta década *A Grammar of Contemporary English* (1972), de Quirk, Greenbaun, Leech y Svartvik, obra en la que se incluye una breve, pero reflexiva,

³⁷ The first problem before us is to identify idiomativity in language structure, to see what idioms are and how they differ from nonidiomatic expressions; the second is to see whether idioms can be “dissolved”, by appropriate analytic devices, into nonidiomatic constituents, and what obstacles stand in the way of such dissolution. (1969: 26)

³⁸ «a complex expression whose meaning cannot be derived from the meaning of its elements» (1969: 26).

³⁹ Finally, to a linguistics that is preoccupied with productivity in the strongest, Chomskyan sense, idiomativity represents a basic theoretical stumbling block; for under the rubric of idiomativity we are concerned with complex structures that can be recognized and analyzed but not naturally generated by any explicit machinery so far proposed. (1969: 22)

discusión sobre expresiones como *How do you do?*, *Why get so upset*, *May the best man win*, o *Least said*. Los autores concluyen que todas ellas son formulaicas debido a sus anomalías gramaticales y a su naturaleza conversacional (Pawley 2007: 12). Aparece en este período el anteriormente mencionado artículo «Regularity and Idiomaticity in Grammatical Constructions: The Case of Let Alone» (1988), del lingüista estadounidense Charles J. Fillmore y sus colaboradores, y que resulta sustancial para el estudio de lo idiomático y prefabricado, en concreto, para la teoría de la gramática de construcciones. En palabras de Hernández Muñoz, este artículo «sienta las bases de una nueva unidad lingüística: la construcción como unidad básica del lenguaje» (2019: 13).

En «Regularity and Idiomaticity in Grammatical Constructions: The Case of Let Alone», Fillmore *et al.* realizan una investigación sobre la sintaxis, la semántica y la pragmática de la construcción *let alone*. Los autores defienden la prevalencia de la noción de construcción frente a las teorías basadas en las reglas de la gramática generativa:

Este artículo aboga por un enfoque de la gramática que difiere de la mayoría de los enfoques actuales en varios aspectos. La afirmación general es que las unidades propias de una gramática son más similares a la noción de construcción en las gramáticas tradicionales y pedagógicas que a la de regla en la mayoría de las versiones de la gramática generativa. (1988: 501)⁴⁰

Fillmore *et al.* afirman que las construcciones aportan, además de información sintáctica, información léxica, semántica y pragmática, asegurando que esta última fue desestimada por la tradición generativista: «ningún enfoque formal actual de la gramática permite una interpretación pragmática directa de las estructuras sintácticas» (1988: 502)^{41y42}. Hernández Muñoz afirma que «estas construcciones no solo vehiculan información sintáctica, sino también léxica, semántica y pragmática» (2019: 13).

⁴⁰This paper advocates an approach to grammar that differs from most current approaches in several ways. The overarching claim is that the proper units of a grammar are more similar to the notion of construction in traditional and pedagogical grammars than to that of rule in most versions of generative grammar. (1988: 501)

⁴¹«no current formal approach to grammar countenances direct pragmatic interpretation of syntactic structures» (1988: 502).

⁴²Para finales de los setenta, y estimulados por el trabajo previo realizado por filósofos del lenguaje como Austin y Searle, una gran cantidad de gramáticos comenzó a investigar sobre la semántica y la pragmática de los actos del habla, poniendo el foco en las construcciones convencionales del lenguaje (Pawley 2007: 12).

En cuanto al reconocido valor idiomático de las construcciones, Fillmore *et al.* sostienen la idea de que una construcción es idiomática en cuanto a que esta puede tener un significado único, es decir, una construcción puede aportar una idea o información sinigual que no es posible deducir mediante el análisis individual de los elementos que la conforman:

Las construcciones pueden ser idiomáticas en el sentido de que una construcción grande puede especificar una semántica (y / o pragmática) que es distinta de lo que podría calcularse a partir de la semántica asociada del conjunto de construcciones más pequeñas que podrían usarse para construir el mismo objeto morfosintáctico. (1988: 501)⁴³

Fillmore *et al.* concluyen que no es posible dar una respuesta a las irregularidades del lenguaje simplemente con unas reglas gramaticales generales y con una lista de estructuras prefijadas. Es necesaria una investigación más profunda del fenómeno de lo idiomático:

... en la construcción de la gramática, se necesita más que un sistema de reglas gramaticales generales y un léxico de palabras y frases fijas. Los procesos lingüísticos que se consideran irregulares no pueden explicarse mediante la creación de listas de excepciones: la esfera de la idiomatidad en una lengua incluye una gran cantidad que es productiva, altamente estructurada y digna de una investigación gramatical seria. (1988: 534)⁴⁴

4.2. Identificación del lenguaje formulaico

Numerosas investigaciones se han embarcado en el proceso de identificación de LF en textos. Uno de los estudios más famosos es «The Idiom Principle and the Open Choice Principle» (2000), de Britt Erman y Beatrice Warren. Mediante el

⁴³ constructions may be idiomatic in the sense that a large construction may specify a semantics (and/or pragmatics) that is distinct from what might be calculated from the associated semantics of the set of smaller constructions that could be used to build the same morphosyntactic object. (1988: 501)

⁴⁴ ... in the construction of grammar, more is needed than a system of general grammatical rules and a lexicon of fixed words and phrases. Those linguistic processes that are thought of as irregular cannot be accounted for by constructing lists of exceptions: the realm of idiomatity in a language includes a great deal that is productive, highly structured, and worthy of serious grammatical investigation. (534)

análisis de un corpus de textos, Erman y Warren demuestran que entre el 52% y el 58% de los textos están formados por LF. Erman y Warren afirman:

Otro aporte importante es la revelación de que existe una gran cantidad de lenguaje prefabricado tanto en textos hablados como escritos (en la mitad de los textos como promedio), lo que imposibilita considerar los modismos y otras combinaciones de múltiples palabras como fenómenos marginales. (2000: 29)⁴⁵

Los autores proponen el término *prefab* para este tipo lenguaje:

Un prefabricado es una combinación de al menos dos palabras preferidas por los hablantes nativos en lugar de una combinación alternativa que podría haber sido equivalente si no hubiera sido convencionalización. (Erman y Warren 2000: 31)⁴⁶

Aun así, la identificación de LF es un arduo y difícil proceso por la cantidad de unidades que pueden identificarse y por la falta de un marco lingüístico que sea referente a la hora de establecer qué es y qué no es una construcción. En *Fundamentals of Formulaic Language: An introduction* (2015), Wood expone la variedad de métodos que históricamente han sido utilizados para identificar construcciones, resaltando dos categorías principales: métodos basados en la medición empírica y métodos basados en un criterio (2015: 20). La decisión más apropiada sería utilizar una combinación de ambos métodos, sin olvidar que «independientemente de las medidas utilizadas para determinar lo formulaico, la certeza absoluta es esquivia» (Wood 2015: 32)⁴⁷.

4.2.1. Métodos empíricos

Dentro de esta categoría podemos encontrar el método estadístico. Este método se basa en el análisis masivo de corpus. Con el objetivo de que el análisis de corpus sea lo más productivo posible, los investigadores establecen dos parámetros principales para la identificación de secuencias: la frecuencia y la

⁴⁵ Another important contribution is the revelation that there is a large amount of prefabricated language in both spoken and written texts (on average on half of the texts), which makes it impossible to consider idioms and other multi-word combinations as marginal phenomena. (2000: 29)

⁴⁶ A *prefab* is a combination of at least two words favored by native speakers in preference to an alternative combination which could have been equivalent had there been no conventionalization. (Erman y Warren 2000: 31).

⁴⁷ «regardless of the measures used to determine formulaicity, absolute certainty is elusive» (Wood 2015: 32).

longitud mínima de las construcciones. Por un lado, las secuencias de LF son generalmente recurrentes, así que una secuencia de palabras que se repita asiduamente es considerada una construcción. La frecuencia de la construcción puede oscilar entre 10 y 40 casos por millón de palabras (Woods 2015: 20-1). Por otro, es necesario determinar la cantidad mínima de palabras que forman una secuencia —por lo general, suele establecerse en tres o cuatro palabras—. Una vez establecidos estos parámetros, se procede al escaneo y análisis del corpus. Este método resulta especialmente útil para el análisis de grandes corpus pertenecientes a disciplinas concretas, compuestos por, incluso, millones de palabras. Así, por ejemplo, puede determinarse qué construcciones son prototípicas en biología, derecho, historia, etc.

Sin embargo, estos análisis empíricos tienen una serie de inconvenientes inherentes a su propia metodología. Por un lado, no resultan productivos con pequeñas bases de datos. Los textos pueden contener LF, pero este no será detectado ya que no aparecerá el número de veces necesario. Por otro, requieren la eliminación de muchas combinaciones de palabras que por su frecuencia son erróneamente identificadas como LF. En *Formulaic Language and the Lexicon* (2002), Wray remarca los problemas de los métodos empíricos:

A primera vista, parece completamente razonable realizar búsquedas con ordenadores para identificar cadenas comunes de palabras y establecer un cierto umbral de frecuencia como criterio para llamar «fórmula» a una cadena... [sin embargo] pueden identificarse problemas relacionados con los procedimientos basados en el recuento de frecuencia. En primer lugar, mediante los corpus no es posible capturar la verdadera distribución de ciertos tipos de secuencias formulaicas... El segundo problema es que las herramientas utilizadas en el análisis de corpus no ofrecen más ayuda que los jueces nativos para decidir dónde recaen los límites entre las secuencias formulaicas. (2002: 25-8)⁴⁸

Por su parte, Wood argumenta que la frecuencia no debe ser el único factor que tener en cuenta a la hora de catalogar una combinación de palabras como una construcción: «es necesario que una cadena sea más que *frecuente*, debe tener

⁴⁸ It seems, on the surface, entirely reasonable to use computer searches to identify common strings of words, and to establish a certain frequency threshold as the criterion for calling a string “formulaic” ... [however] problems regarding the procedures of frequency counts can be identified. Firstly, corpora are unable to capture the true distribution of certain kinds of formulaic sequences ... The second problem is that the tools used in corpus analysis are no more able to help decide where the boundaries between formulaic sequences fall than native speaker judges are. (2002: 25-8)

un significado o función unitario, y quizá también una forma particular de ser mentalmente almacenada, recuperada o producida» (2015: 20)⁴⁹. Asimismo, los famosos *chunks* también pueden resultar problemáticos. Secuencias como, *por un lado, en gran medida, o ¿qué tal estás?*, son secuencias de carácter interdisciplinar, y, en consecuencia, pueden o no aparecer con frecuencia en corpus de una temática específica. Con todo, y de acuerdo con Wood, los *chunks* son claramente LF: los elementos que las conforman tienden a «pegarse», y, además, tienen unos significados o funciones concretas (2015: 21). Una posible solución para identificar este tipo de construcciones mediante un método empírico es a través de procedimientos como la información mutua —*mutual information*—. En lingüística, el procedimiento de información mutua determina cuál es la probabilidad de que las palabras que constituyen una construcción aparezcan juntas. Una alta puntuación en *MI* indicaría una mayor probabilidad de coocurrencia, que, junto con unos datos de frecuencia notables, evidenciaría la existencia de una construcción o secuencia formulaica (Woods 2015: 21).

En cuanto a las propuestas más recientes para la identificación de fórmulas mediante métodos empíricos, destaca el original estudio realizado por C. C. Shei en 2008 «Discovering the hidden treasure on the internet: Using Google to uncover the veil of phraseology». Shei propone que tanto profesores como alumnos utilicen internet como si se tratase de un inmenso corpus, usando motores de búsqueda como el de Google para identificar, investigar y enseñar el LF. Shei sostiene que, si introducimos una secuencia de palabras en Google y examinamos sus resultados, podemos obtener valiosa información sobre su frecuencia, forma, variabilidad y funciones (Woods 2015: 22). Por último, desde la psicolingüística y la fonología también se han desarrollado métodos con inclinación empírica. Los más reseñables en la psicolingüística han sido los basados en el rastreo de los movimientos oculares durante la lectura y los basados en la medición de la latencia —velocidad de reacción durante la lectura—. Por parte de la fonología, se han desarrollado métodos de identificación basados en el análisis de las características prosódicas de las construcciones, como carencia de pausas y titubeos, ritmo prominentemente alto y restricciones en la ubicación del acento. Igualmente, estos últimos métodos tampoco resultan cien por cien satisfactorios para la identificación de LF. De acuerdo con Wood, tanto los métodos psicolingüísticos como los fonológicos proporcionan «una idea limitada o

⁴⁹ «it is necessary for a string to be more than just frequent, it needs to have unitary meaning or function, and perhaps a particular way of being mentally stored, retrieved, or produced as well» (2015: 20).

parcial sobre lo formulaico» (2015: 23)⁵⁰. Ambos métodos están «limitados únicamente al análisis del lenguaje oral. Además, [ambos solo se relacionan] con las fórmulas utilizadas por un hablante en particular, y el análisis está limitado por la calidad de los audios grabados» (2015: Woods 23)⁵¹.

4.2.2. Otros criterios

Por lo que concierne a los métodos basados en un criterio, Wood destaca que estos pueden ser una alternativa a los métodos empíricos (2015: 23). Existen diversos métodos dentro de esta categoría, pero el más desarrollado es el juicio del hablante nativo. Este método se fundamenta en la aplicación del criterio lingüístico por parte de unos «jueces» para determinar la existencia de LF. Dicho de otra forma, una serie de, generalmente, expertos en lingüística, determinan gracias a su criterio si existe LF o no. Este método es especialmente útil a la hora de identificar secuencias en corpus de lengua oral, ya que el criterio humano es superior al tecnológico a la hora de analizar aspectos como la entonación o los patrones de pausas.

No obstante, Wray apunta una serie de inconvenientes al utilizar este método: su uso debe de estar restringido a pequeñas bases de datos, las valoraciones pueden terminar siendo inconsistentes debido a la fatiga, pueden existir discrepancias entre los evaluadores y estos pueden llegar a tener dudas sobre qué identificar; y la intuición como fórmula para identificar estructuras puede aparecer debido a la falta de conocimiento sobre el tema (2002: 23). En cuanto al problema relacionado con el uso de la intuición, Wray afirma:

Una objeción para la intuición como forma de juzgar la estructura lingüística proviene de la investigación de corpus. Antes de la aparición de la tecnología necesaria para la búsqueda en grandes corpus, generalmente se asumía que nuestras intuiciones sobre el lenguaje eran esencialmente precisas, por lo que parecía haber poca diferencia entre encontrar un ejemplo ilustrativo en un texto real o inventarlo. Sin embargo, la investigación de corpus ha revelado que «la intuición humana sobre el lenguaje es muy específica y y para nada es una buena guía de lo que realmente sucede cuando, de hecho, unas mismas personas usan la lengua» (2002: 21).⁵²

⁵⁰ «a limited or partial sense of formulaicity» (2015: 23).

⁵¹ «limited to analysis of spoken language only. As well, [both only relate] to formulas used by a particular speaker, and analysis is limited by the quality of audio data recorded» (Woods 2015: 23).

⁵² A quite different objection to intuition as a way of judging linguistic structure comes from corpus

De acuerdo con Wray, la intuición no resulta un método completamente fiable ni siquiera para los hablantes nativos, ya que estos tienden a categorizar sin rigor científico y guiados por su intuición un gran número de expresiones como construcciones que realmente no lo son (2002: 22). En su lugar, Wray propone estas propiedades para facilitar la identificación de secuencias formulaicas:

1. *Estructura*: en muchas ocasiones, las construcciones comienzan con preposiciones, artículos, conjunciones, pronombres o marcadores discursivos (2002: 31-3). Por ejemplo, *para finales de semana/mes/año*.
2. *Composicionalidad*: estructura interna de la construcción. La construcción no debe ser necesariamente gramatical o semánticamente lógica. (2002: 33-4)
3. *Fijación —fixedness—*: tendencia por la cual las construcciones prefabricadas son invariables en su forma, aunque existen muchas con huecos —*slots*— que pueden ser ocupados por distintas palabras (2002: 34)
4. *Forma fonológica*: las secuencias pueden tener una prosodia particular con una entonación propia (2002: 35)
5. *Fluidez, acento y articulación*: las secuencias suelen carecer de pausas internas, son difíciles de acentuar de distintas maneras y su articulación es más imprecisa. (2002: 35-9)

Similar a lo propuesto por Wray, una posible solución para solventar los problemas del método del juicio del hablante nativo es utilizar una lista de comprobación —*checklist*— con una función orientativa. Este *checklist* debe contener características prototípicas del LF. Los jueces examinan el *checklist*, para después, aplicarlo en el proceso de identificación de secuencias. Existen diversos *checklist*, algunos más concretos y otros más generales. Conforme con Wood, los más importantes han sido los elaboradas por los siguientes autores: Coulmas (1979), una de las primeras listas en usarse; Peters (1983), lista para identificar el LF en la adquisición del lenguaje en niños; Wood (2006, 2009, 2010), lista para

research. Before the advent of the technology for searching large corpora, it was generally assumed that our intuitions about language were basically accurate, so it seemed to make little difference whether you found an illustrative example in real text or made one up. However, corpus research has revealed that “human intuition about language is highly specific, and not at all a good guide to what actually happens when the same people actually use the language”. (2002: 21)

identificar el LF en el aprendizaje y fluidez de en segunda lengua; y Wray y Namba (2003), lista aplicable para niños y adultos, nativos o no nativos (2015: 25).

4.3. Taxonomía

El LF abarca una gran cantidad de categorías. En gran medida, esto es debido a la variedad de disciplinas que han estudiado este fenómeno. Cada disciplina ha aportado un enfoque distinto, y, por ende, la clasificación de las construcciones ha variado a lo largo del tiempo. Realizar, por lo tanto, una categorización meticulosa del LF es una tarea muy complicada. En este sentido, Wray afirma lo siguiente sobre la taxonomía del LF:

Las taxonomías son atractivas porque prometen una clasificación ordenada. Sin embargo, para que sean útiles o teóricamente plausibles, deben ser coherentes internamente, con algún tipo de dependencia de principios entre las clases que diferencian. De lo contrario, no son más una glorificación de listas. Además, debe haber alguna forma clara de mantener las categorías separadas. (2002: 47)⁵³

En «The functions of formulaic language: an integrated model» (2000), Wray y Perkins afirman que han llegado a contabilizarse más de cuarenta tipos distintos de construcciones (2000: 3) (Anexo 1). Ambos autores apuntan que esta diversidad de términos responde a la multitud de formas por las que se ha caracterizado al lenguaje formulaico: «La naturaleza polifacética del lenguaje formulaico es evidente por la variedad de formas en que se ha caracterizado: según su forma, función, propiedades semánticas, sintácticas y léxicas, y su relación con el lenguaje novedoso (analítico)» (2000: 3)⁵⁴. No obstante, Wray afirma que, generalmente, las taxonomías del LF han tenido en consideración especial cuatro características concretas: forma, función, significado y procedencia (2002: 47). De acuerdo con Wray, el problema que conlleva considerar únicamente estas cuatro características es la superposición, es decir, producir taxonomías difusas

⁵³ Taxonomies are attractive because they promise a neat categorization. However, in order to be either useful or theoretically plausible, they must be internally consistent, with some sort of principled dependency between the classes they differentiate. Otherwise, they are no more than glorified lists. Also, there must be some clear way of keeping categories separate (2002: 47)

⁵⁴ «The multi-faceted nature of formulaic language is evident from the variety of ways in which it has been characterised: according to its form, function, semantic, syntactic and lexical properties, and its relationship with novel (analytic) language» (2000: 3).

que tiendan a entremezclarse: «el problema es que estas características se superponen, lo que complica más la situación» (2002: 47)⁵⁵.

A pesar de la magnitud taxonómica mencionada y de los problemas derivados por intentar establecer delimitaciones rígidas, existen unas categorías particularmente interesantes para los investigadores sobre el LF. Conforme con Wood, estas categorías son: colocaciones, modismos, grupos sintagmáticos (*lexical phrases*), paquetes léxicos (*lexical bundles*), metáforas, proverbios, verbos proposicionales (*phrasal verbs*), n-gramas, *congrams*⁵⁶ y compuestos (2015: 37). Wood concluye asumiendo que el LF es un fenómeno poco unitario cuya clasificación resulta complicada y, en ocasiones, arbitraria:

A partir de esta descripción general de las categorías del lenguaje formulaico, puede resultar sorprendente cuantos tipos existen y lo variados que son. El fenómeno que tratamos no es de ninguna manera unitario, y las clasificaciones y taxonomías son algo imprecisas o débiles; las distinciones entre, por ejemplo, una colocación y un modismo son difusas, y además parece que determinados investigadores han compuesto de forma algo arbitraria sus propios conjuntos de descripciones, clasificaciones y definiciones de los diversos tipos de LF. (2015: 50)⁵⁷

⁵⁵ «The problem is that these features overlap, causing some muddying of the waters» (2002: 47).

⁵⁶ En inglés, un *congram* son todas las permutaciones posicionales que admite un conjunto de dos o más palabras.

⁵⁷ From this general overview of categories of formulaic language, it can be surprising how many and varied the types are. The phenomenon we are dealing with is by no means unitary, and the classifications and the taxonomies are somewhat leaky or slippery—the distinctions between, for example, a collocation and an idiom are blurry, and it also appears that particular researchers have somewhat arbitrarily composed their own sets of descriptions and classifications and definitions of the various types. (2015: 50)

5. Funciones de las fórmulas: un enfoque etnolingüístico

Como ha sido comentado anteriormente, una de las características del LF es su idomaticidad. Muchas secuencias de lenguaje prefabricado tienen un significado concreto que no es posible transmitir mediante otras palabras. En este sentido, entra en juego las funciones comunicativas que cumplen estas fórmulas y su componente sociopragmático. Moreno Teva apunta a este respecto:

El lenguaje formulaico presenta el léxico como un reflejo directo de la manera en que el lenguaje opera para el individuo en su comunidad o comunidades de habla ... el léxico solamente admite aquellas unidades que, grandes o pequeñas, han sido identificadas como comunicativamente útiles por la experiencia directa y, en algunos casos, cuando se ha producido una alta ocurrencia observada por el individuo ... La convencionalización de unidades lingüísticas viene determinada normalmente por la totalidad de una comunidad de habla, aunque algunas son empleadas solamente por determinados grupos sociales, dentro de ciertos campos o comunidades científicas (en economía, biología o en el campo jurídico, por ejemplo) o incluso pueden ser características de un individuo en concreto. (2012: 20-1)

En su tesis, *The Role of Formulaic Language in the Creation of Grammar* (2011), Stukova apunta la idea de que el LF cumple, por un lado, una función facilitadora en términos de esfuerzo y procesamiento del lenguaje —facilitando así la fluidez—, y por otro, una función solidaria entre el emisor y el destinatario debido a su indexicalidad:

El lenguaje formulaico es una estrategia en los idiomas para lograr los objetivos simultáneos de eficiencia en la expresión y de facilidad en el procesamiento, además del importante objetivo de construir una solidaridad entre un hablante y un oyente

mediante el uso de frases conocidas para expresar un mensaje específico. Lo prefabricado tiene, en este sentido, una importante indexicalidad social. (2011: 137)⁵⁸

En relación con esta función solidaria o social del LF, recientemente Olímpio de Oliveira Silva expone en un interesante artículo la aplicación de las investigaciones realizadas desde la etología humana a la lingüística, concretamente al campo de la fraseología⁵⁹. La autora afirma que la ciencia de la etología humana no ha sido lo suficientemente aprovechada por los lingüistas, más allá de la creación en los años setenta de una etnolingüística que pretendía dar respuestas al porqué de nuestras elecciones a la hora de comunicarnos (2015: 157-8). Olímpio de Oliveira Silva propone establecer un punto de contacto entre ambas disciplinas. A partir de esta novedosa propuesta, pueden obtenerse interesantes ideas sobre las funciones sociales que cumplen las fórmulas rutinarias en la comunicación. Olímpio de Oliveira Silva define las fórmulas rutinarias como «fórmulas de la interacción social habituales y estereotipadas que cumplen funciones específicas en situaciones predecibles, rutinarias y, hasta cierto punto, ritualizadas» (*apud* Olímpio De Oliveira Silva 2015: 152). Como puede observarse, estas fórmulas rutinarias son, esencialmente, lo que desde la lingüística inglesa se ha estudiado en los últimos años como LF, es decir, estructuras prefabricadas y altamente automatizadas. De acuerdo con Olímpio de Oliveira Silva, lo realmente interesante al aplicar la etnología humana al ámbito lingüístico es examinar la influencia que tienen los factores ambientales y genéticos en los comportamientos que adoptamos los humanos durante la comunicación (2015: 154). Por un lado, los factores genéticos están condicionados por un sistema universal de reglas —filogenético e innato—, y por otro, los factores ambientales están condicionados por una gramática cultural —contextual y social— (2015: 156). El propósito es, en definitiva, comprender qué estructuras y patrones lingüísticos son intrínsecos al ser humano, y a su vez, identificar cuáles son aprendidos o adquiridos.

⁵⁸ Formulaic language is one strategy that languages have to accomplish the simultaneous goals of efficiency of expression and ease of processing, in addition to the important aim of building of solidarity between a speaker and a hearer by using familiar phrasing to express a specific message. Prefabs have, in this sense, an important social indexicality. (2011: 137)

⁵⁹ A pesar de que Olímpio se centra en la aplicación de la etnolingüística a la fraseología, su artículo posee un gran valor para el estudio de las funciones del LF. Como se ha comentado, Wood apunta que la fraseología es el término utilizado para estudiar unas categorías concretas de LF.

En cuanto a la influencia de los factores innatos en la comunicación, esta puede observarse en las metáforas que oponen *claro* y *oscuro*, recurrentes, además, en una gran diversidad de idiomas:

La contribución de la Etología Humana al entendimiento de las metáforas puede ilustrarse a partir del enfoque según el cual las imágenes relacionadas con la oposición claro-oscuro (*personalidad radiante, ideas claras o tener algo claro frente a a oscuras* —con la acepción de ‘sin conocimiento de algo, sin comprender lo que se oye o se lee’— o *personaje oscuro*) podrían explicarse en función de la predilección que tienen los humanos por el modo de vida diurno, el tiempo en el que se desarrollan las actividades, que se relaciona con la seguridad, en oposición con el temor que infunde la noche. Estas metáforas no vendrían dadas, al menos desde un principio, por cuestiones socioculturales (como puede ser el racismo). (Olímpio de Oliveira Silva 2015: 157)

Olímpio de Oliveira Silva menciona en su artículo cuatro tesis etnolingüísticas interrelacionadas entre sí y relevantes para investigar los comportamientos en la comunicación humana.

La primera teoría alude a la idea de que determinados actos verbales pueden ser el equivalente a actos no verbales. En otras palabras, el ser humano transmite unas intenciones a través del lenguaje. Lo relevante respecto a las construcciones, a las unidades fraseológicas, y, en definitiva, al LF, es que a partir de esta premisa puede aplicarse la reconocida teoría de que la forma, el significado y el uso de lenguaje prefabricado está notablemente relacionado con la comunicación no verbal (Olímpio de Oliveira Silva 2015: 159). *Cruz y raya* o [quedarse] *de brazos cruzados* son ejemplos de construcciones que generalmente son acompañadas de una comunicación no verbal. De esta manera, podría investigarse hasta qué punto el uso de un lenguaje prefabricado responde a aspectos intrínsecos del ser humano o a aspectos ambientales o contextuales.

La segunda teoría propuesta guarda relación con la primera, ya que sostiene que «el habla, al sustituir a la acción, permite refrenar ciertas actitudes más instintivas, como la agresividad, y también transmutarlas» (Olímpio de Oliveira Silva 2015: 161). Esto significa que el ser humano tiene la capacidad mediante el lenguaje de transmitir unas intenciones que, de otra manera, causarían un conflicto entre los interlocutores. Olímpio de Oliveira Silva expone que desde la etología se ha estimado que la agresividad es una característica innata, es decir, universal a todos los humanos (2015: 162). Se justifica así la existencia de

un lenguaje prefabricado constituido por fórmulas rutinarias que es utilizado como una herramienta lingüística y social para transmitir ideas o actitudes, en mayor o menor grado, propensas al enfrentamiento. Este tipo de LF o prefabricado es común en multitud de lenguas. Algunos ejemplos en la lengua española son los siguientes:

*No me hagas hablar
Para que te enteres
Se va a enterar
Tengamos la fiesta en paz
¡Tú mismo!
Tu verás
Ya nos veremos las caras
Ya te acordarás
Ya te apañaré
Ya te guardarás*

Fórmulas de promesa y amenaza (Olímpio de Oliveira Silva 2015: 161).

La tercera tesis está relacionada con la segunda, y trata sobre el alto grado de ritualización del comportamiento hablado. Olímpio de Oliveira Silva señala al respecto que «un ritual viene a ser un patrón de comportamiento al servicio de la función comunicativa y que este puede sufrir cambios a favor de esa función» (2015: 162-3). Los rituales están al servicio de la intención comunicativa que subyace. Se deduce, por lo tanto, la existencia de una serie de reglas en los rituales para conseguir los propósitos comunicativos. En relación con el lenguaje prefabricado, esta ritualización conecta con la idea de que las fórmulas tienen un carácter predecible, son rutinarias, están sujetas a convenciones y cumplen una función comunicativa (Olímpio de Oliveira Silva 2015: 163-4). Además, estas fórmulas son universales, ya que a pesar de que los factores culturales y sociopragmáticos puedan alterar en distinto grado sus formas y su utilización, estas han sido encontradas en todas las lenguas estudiadas. Olímpio de Oliveira Silva señala como un claro ejemplo comunicativo ritualizado las distintas fórmulas para pedir disculpas:

Así, por ejemplo, entre las estrategias lingüísticas que configuran un acto de habla específico y altamente ritualizado, como el de disculparse, se encuentran diferentes tipos de fórmulas que ayudan a resaltar su carácter ritual. Existen, pues, diferentes fórmulas de disculpa (*lo siento (mucho), ¡cómo lo siento!, ha sido sin querer, con perdón*); de pseudodisculpa (*perdone que le moleste, perdone que le interrumpa*) y fórmulas de réplica —en este caso, para responder a la petición de disculpas— (*no*

ha sido nada, no hay de qué, no pasa nada, no tiene importancia). (Olímpio de Oliveira Silva 2015: 164)

Por último, la cuarta tesis se fundamenta en la idea del lenguaje como mecanismo armonizador de la vida en sociedad. Ya hemos visto en las teorías anteriores como el habla tiene un carácter ritual y puede servir para distanciarse de lo instintivo. Pues bien, estas premisas confluyen en establecer, finalmente, el principio de que el lenguaje humano utiliza fórmulas como mecanismo para conciliar la convivencia, evitar enfrentamientos y conservar las relaciones personales. Al utilizar un lenguaje prefabricado, predecible y automatizado, reducimos la posibilidad de crear malentendidos, producir ambigüedades o provocar confrontaciones (Olímpio de Oliveira Silva 2015: 166). De acuerdo con Olímpio de Oliveira Silva, esta teoría sirve para identificar una serie de fórmulas sin categoría específica, las cuales pueden asociarse con la solidaridad o el ánimo:

<p><i>En paz Haya paz (Y) Ya está</i></p>	<p><i>Eso pasa / ocurre en las mejores familias Más se perdió en Cuba No hay para tanto No pasa nada ¡Qué más da! ¡Ya será menos!</i></p>
---	---

Fórmulas de solidaridad y ánimo (Olímpio de Oliveira Silva 2015: 165-6).

Asimismo, otras fórmulas que pueden aportar armonización a la vida en grupo son las que Olímpio de Oliveira Silva menciona como fórmulas de información. Estas fórmulas están relacionadas con la petición de favores, objetos, ayudas, etc. Con el fin de mantener una relación cordial, el individuo que necesite pedir algo a otro deberá hacerlo de forma respetuosa y sin presionarlo. Si no existe una relación estrecha entre los interlocutores, la falta de esta habilidad comunicativa puede conllevar al rechazo o a la incomunicación. Estamos entrando, por lo tanto, en el terreno del conocido fenómeno de la cortesía verbal⁶⁰. Olímpio de Oliveira Silva señala en su artículo algunas de las construcciones más utilizadas en español para estos fines:

⁶⁰ En lingüística, se entiende por cortesía el conjunto de estrategias conversacionales destinadas a evitar o mitigar las tensiones que aparecen cuando el hablante se enfrenta a un conflicto creado entre sus objetivos y los del destinatario. (*Diccionario de términos clave de ELE*)

¿Haces el favor?
 ¿Me haces el favor?
 ¿Me permite (s)?
 ¿(No) le importa/importaría que + subj ? // ¿(no) le importa si + ind ? // si no le importa + ind // si no le importa que + sub// ¿No le importa + subj ? // no le importa + subj , ¿verdad?
 ¿Sería (usted) tan amable de + infin o prop con QUE?
 // si es (usted) tan amable de + infin o prop con QUE
 // si fuera (usted) tan amable de + infin o prop con QUE
 Si no le sirve de molestia
 ¿Tendría (usted) la amabilidad de + infin ? // tenga (usted) la amabilidad de + infin Tenga la bondad // si tiene la bondad // ¿tendría la bondad? // ¿tiene la bondad?

Fórmulas para peticiones o favores (Olimpio de Oliveira Silva 166).

Como podemos observar, la aplicación de un enfoque interdisciplinar para el estudio del lenguaje prefabricado proporciona una valiosa información sobre sus funciones innatas. Debe asumirse pues, que no es posible dar una explicación total al fenómeno de lo idiomático desde perspectivas restringidas únicamente a factores culturales o sociales. El porqué del uso de determinadas fórmulas o construcciones atiende también a razones genéticas y evolutivas. Mediante la aplicación de un enfoque etnológico y lingüístico, Olímpio de Oliveira Silva demuestra la naturaleza del uso de una serie de fórmulas comunicativas de carácter ritual, cuya característica principal es el avenimiento de la vida en sociedad. Esta función social del lenguaje es, en el fondo, la reconocida dimensión *agente social* recogida tanto en el *MCER* como en el *Plan Curricular del Instituto Cervantes*. Esta dimensión del hablante establece que el aprendiente necesitará utilizar la lengua como herramienta vehicular para cumplir satisfactoriamente las tareas propias del individuo inmerso en una sociedad:

La dimensión del *uso social* viene dada por la naturaleza misma de la lengua como instrumento de comunicación ... Los treinta años transcurridos desde la definición del nivel umbral hasta la publicación del *MCER* en 2001 han permitido ahondar en la reflexión sobre las implicaciones de la idea misma de uso social de la lengua ... El alumno como *agente social*, ... ha de conocer los elementos que constituyen el sistema de la lengua y ser capaz de desenvolverse en las situaciones habituales de comunicación que se dan en la interacción social (*Plan Curricular del Instituto Cervantes*).

6. Adquisición del lenguaje formulaico

6.1. Adquisición en niños (L1)

El lenguaje utilizado por los niños proporciona una valiosa información para comprender mejor el funcionamiento del lenguaje prefabricado y su importancia en el desempeño comunicativo en edades tempranas. La adquisición del lenguaje puede interpretarse como un proceso de «captura de piezas». En un primer período que abarca aproximadamente los veinticuatro primeros meses, los niños almacenan unidades lingüísticas procedentes de su entorno. Una vez superada esta fase, configuran internamente estas unidades y comienzan a manejarlas y modificarlas. Esta última fase comienza a los dos años y se extiende aproximadamente hasta los seis⁶¹. Wray define las dos etapas mencionadas anteriormente como la fase de aprendizaje de unidades —*item-learning*— y la fase de aprendizaje del sistema —*system-learning*— respectivamente (2002: 106), atisbándose en esta última el proceso de segmentación y adecuación gramatical:

las palabras compuestas y las frases nominales que involucran varios morfemas son comúnmente aprendidas por los niños como conjuntos sin analizar. *Taza de té* se puede aprender como una unidad y, cuando se ha aprendido el morfema plural, se puede pluralizar como *tazas de té*. (*apud* Wray 2002: 106)⁶²

⁶¹ Las etapas en la adquisición de la lengua son generalmente divididas de la siguiente manera: etapa prelingüística (de 0 a 12 meses); etapa lingüística (a partir de los 12 meses); período holofrástico (desde los 12 meses hasta los 2 años); período de expansión sintáctica (desde los 2 años hasta los 6 años).

⁶² compound words and noun phrases involving several morphemes are commonly learned by children as unanalyzed wholes. Cup of tea may be learned as one unit and, when the plural morpheme has been learned, it may be pluralized as cup of teas. (*apud* Wray 2002: 106)

Desde la teoría de la gramática de construcciones también se apoya esta misma hoja de ruta en el proceso de adquisición del lenguaje. Diessel defiende que el proceso comienza con la adquisición de fórmulas. Posteriormente, estas fórmulas sufren una segmentación en unidades más complejas. En última instancia, pasan a formar parte del entramado general de construcciones. En palabras de Diessel,

Se ha argumentado que el desarrollo gramatical de los niños comienza con unas fórmulas léxicas específicas que gradualmente se van descomponiendo, elaborando así unidades más complejas y esquemáticas... Lo que los niños finalmente aprenden es una red de construcciones que se basa inmediatamente en su experiencia lingüística. (2013: 13)⁶³

En un primer momento, podría considerarse que el lenguaje prefabricado no debería aparecer hasta un punto avanzado en el proceso de adquisición. Las secuencias parecen demasiado complejas para un niño de entre doce y veinticuatro meses. Todo lo contrario. Es normal que durante este período los niños reproduzcan un lenguaje prefabricado. En *The Units of Language Acquisition* (1983), Peters investiga el papel de este lenguaje en los niños. La autora comenta:

El niño no está expuesto a un diccionario de morfemas, sino más bien un flujo intermitente de sonidos que contienen *chunks*, a menudo más largos que una sola palabra, y que se repiten con una frecuencia variable. A partir de este influjo de significado y estructura desconocidos el niño debe intentar capturar algunas piezas con el fin de determinar su significado y preservarlas para un uso futuro. (1983: 5)⁶⁴

Con el fin de comprender mejor esta, aparente, prematura adquisición de LF, resulta clave explicar la percepción que tienen los niños sobre lo que es una *unidad*. Es ampliamente aceptado, que durante la *item-learning* fase los niños operan únicamente con unidades aisladas para comunicarse. Desde una

⁶³ It has been argued that grammatical development starts with lexically-specific formulas that children gradually decompose and elaborate to more complex and schematic units ... What children eventually learn is a network of constructions that is immediately grounded in their linguistic experience. (2013: 13)

⁶⁴ It is not a dictionary of morphemes that the child is exposed to, but rather an intermittent stream of speech sounds containing chunks, often longer than a single word, that recur with varying frequency. It is out of this stream of unknown meaning and structure that the child must attempt to capture some pieces in order to determine their meaning and to preserve them for future use. (1983: 5)

perspectiva adulta y analítica, la unidad principal del lenguaje es la palabra. Sin embargo, la idea de unidad es distinta en los niños. Para ellos, una unidad puede llegar a ser una cadena de palabras, ya que su conocimiento lingüístico y sus capacidades cognitivas no están lo suficientemente desarrolladas⁶⁵. Peters afirma al respecto: «Dado que el niño no conoce el idioma, no es razonable asumir que las primeras unidades que extraiga coincidirán exactamente con las palabras y morfemas del sistema» (1983: 5)^{66 y 67}. Wray apunta, además, que la comprensión analítica no es una prioridad en estas edades:

Mucho de lo que el niño lingüísticamente ingenuo oye cada día no debe y no puede tener un sentido particular dada la falta de conocimiento del mundo de un niño de tres años. *¿Tiene sentido?* no está entre las preguntas que preocupan especialmente a los niños pequeños (*apud Wray 2002: 109*)⁶⁸.

No es anómalo pues, que los niños reproduzcan estructuras más complejas de lo esperable, ya que, inconscientemente, catalogan estas cadenas de palabras como unidades mínimas de comunicación⁶⁹: «desde hace mucho tiempo se reconoce que las cadenas bastante largas, que corresponderían a varias palabras

⁶⁵ Esta dicotomía en la concepción de unidad entre adultos y niños «cuestiona el paradigma estándar en la investigación sobre la adquisición del lenguaje que asume que los niños siempre perciben el lenguaje como una disposición de unidades del tamaño de la palabra» (Wray 2002: 111). [«question the standard paradigm in first language acquisition research, that children always perceive language as arrangements of discrete word-size units» (Wray 2002: 111)].

⁶⁶ «Since the child does not know the language, it is unreasonable to assume that the first units she or he extracts will coincide exactly with the words and morphemes of the system» (1983: 5).

⁶⁷ En esta etapa también aparecen las conocidas holofrasas. Una holofrase es una palabra formada a partir de un conglomerado de palabras y con un único significado. Por ejemplo, la holofrase «magua» equivaldría al enunciado *mama dame agua*. Pueden considerarse como los primeros enunciados del niño.

⁶⁸ Much of what the linguistically naive child hears each day must not make particular sense and cannot, given the lack of world knowledge enjoyed by three-year-old. 'Does it make sense?' may not be among the questions that young children particularly concern themselves with. (*apud Wray 2002: 109*)

⁶⁹ Resulta necesario remarcar, que el niño que produce un lenguaje altamente formulaico no se encuentra en un punto más avanzado que aquel que solo produce palabras aisladas. De acuerdo con Wray, «debido a que el tamaño de la unidad puede variar, algunos niños parecerán estar más avanzados que otros cuando, de hecho, todos están sujetos a una «limitación cognitiva que restringe al niño a procesar lo que para él/ella es una única unidad» (2002: 112). [«because the size of the unit can vary, some children will appear to be more advanced than others when, in actual fact, they are all subject to a “cognitive limitation constraining the child to processing (e.g., extracting and producing) what to the child is one unit at a time”» (2002: 112)].

Para determinar el avance en el proceso de adquisición, lo esencial no es contabilizar el número de palabras, sino el número de unidades.

adultas, pueden ser tratadas como una sola unidad por el niño pequeño» (Wray 2002: 105)⁷⁰. Estas cadenas de palabras no sufren ningún proceso de análisis y son adquiridas como un conjunto. Peters sostiene que, como si de palabras se tratase, estas estructuras pasan a formar parte del léxico del niño, por lo que el LF juega un papel primordial en edades tempranas:

... No hay evidencia de que lo que el lingüista reconoce como rutinas no analizadas sea reconocido y almacenado por el niño de manera diferente respecto a los elementos que nosotros llamaríamos palabras sueltas. Por tanto, las rutinas solo están «separadas» del sistema generativo en la medida en que el léxico está separado. Reproducir una rutina no es diferente de reproducir cualquier otra unidad del léxico. (1983: 15)⁷¹

Debido a la carencia de análisis, Wray menciona que resulta habitual que los niños comentan errores al utilizar estas estructuras —malinterpretaciones, descontextualizaciones, reproducciones aleatorias sin un sentido coherente, etc.—, aunque sin graves consecuencias lingüísticas. (2002: 108). Sin embargo, la autora afirma el uso de un lenguaje prefabricado específicamente memorizado que tiene un papel crucial en el aprendizaje: el lenguaje de las rimas, las canciones y las rutinas sociales (108).

En cuanto a las rimas y canciones, los niños las aprenden gracias al ritmo y a las variaciones en el tono. La prioridad del niño es reproducir con la mayor exactitud posible la rima o canción⁷². Es comprensible pues, que el niño no consiga identificar la información que esta pueda transmitir. Wray asume que, por norma general, las rimas y canciones no contienen una información relevante (2002: 109). Sin embargo, la autora también afirma que en ocasiones estas sí aportan una importante información que el niño debe aprender correctamente. Estas son las rimas y canciones sobre el abecedario, los días de la semana, los meses del año, los números, etc. (2002: 109). De acuerdo con Wray, estas rimas y

⁷⁰ «it has long been recognized that quite lengthy strings, which would correspond to several adult words, can be treated as a single unit by the young child» (Wray 2002: 105).

⁷¹ ... there is no evidence that what the linguist recognizes as unanalyzed routines are recognized and stored by the child any differently from items that we would call single words. Hence, routines are only "separate" from the generative system insofar as the lexicon is separate. Reproducing a routine is in no way different from reproducing any other single item from the lexicon. (1983: 15)

⁷² Un dato importante es que algunas palabras aparentemente ya adquiridas pueden ser pronunciadas incorrectamente. Esto significaría que el niño aún no identifica esas palabras como unidades independientes, a pesar de haberlas articulado previamente en otros contextos (Wray 2002: 108).

canciones tienen un carácter nemotécnico, y sirven al niño para almacenar una información útil a la que resultaría difícil acceder de otra manera (2002: 109).

En lo que respecta a las rutinas sociales —o rutinas institucionalizadas— Wray establece que estas son «un conjunto de frases utilizadas en determinadas situaciones para realizar rituales» (2002: 109)⁷³. Normalmente, las rutinas o fórmulas de cortesía son las primeras en enseñarse explícitamente, ya que sirven al niño para empezar a desenvolverse en el entorno adulto. *Por favor, muchas gracias, buenas noches, ¿me puedes/podrías dejar...?* son algunos ejemplos de este tipo de LF. Este lenguaje prefabricado debe ser

recordado como un completo y producido en el momento adecuado para lograr un acto de habla. Emitir una rutina de forma incorrecta puede considerarse un fracaso, incluso cuando la intención es clara, ya que solo las expresiones permitidas se consideran apropiadas dentro de la cultura del hablante. (Wray 2002: 110)⁷⁴

De acuerdo con Wray, el uso de estas fórmulas representa «un reconocimiento de la relación de poder entre el adulto y el niño, ayudando a este último a aprender formas socialmente aceptables de operar en un entorno adulto» (110)⁷⁵.

Por último, Wray afirma que el estudio de los estilos cognitivos también puede dar respuestas a la pregunta del por qué algunos niños producen un lenguaje mucho más prefabricado que otros (Wray 2002: 114)⁷⁶. La autora expone que los niños pueden tener dos estilos cognitivos principales a la hora de comunicarse: referencial y expresivo (2002: 114). El estilo referencial se basa en el uso del lenguaje como herramienta para aludir o etiquetar el mundo más inmediato: objetos, personas, animales, etc. Por el contrario, el estilo expresivo se centra en utilizar el lenguaje como un mecanismo para ejercer algún tipo de control social: reclamar atención, pedir algo, mostrar enfado, etc. Conforme con Wray, el desarrollo de un estilo cognitivo u otro está determinado por factores como la personalidad, la organización neuronal o las primeras experiencias no lingüísticas (2002: 114). Debido a estos diferentes propósitos comunicativos, cada

⁷³ «set of phrases used in certain situations in order to perform rituals» (2002: 109).

⁷⁴ «remembered whole and produced at the right time, in order to achieve a speech act. Saying a routine incorrectly can be counted as a failure, even when the intention is clear, for only the permitted expressions are considered appropriate within the speaker culture. (Wray 2002: 110).

⁷⁵ «an acknowledgement of the power relationship between adult and child and help the child to learn socially acceptable ways of operating in an adult environment» (110).

⁷⁶ En psicología, un estilo cognitivo hace referencia a la forma en la que una persona percibe, procesa y recuerda la información.

estilo cognitivo propicia el uso de un determinado lenguaje. Así, los niños con un estilo cognitivo referencial requieren principalmente el uso de sustantivos, mientras que los niños con un estilo expresivo requieren el uso de cadenas de palabras:

La referencialidad requiere etiquetas individuales para los objetos... La expresividad, por otro lado, requiere de cadenas más largas con una función asociada. (Wray 2002: 114)⁷⁷

El lenguaje que se utiliza para el control social... es probable que se escuche en conjuntos que no se dividen fácilmente; por ejemplo: «¿Te gustaría salir?», «No sé dónde está», «Para con eso». La segmentación de secuencias de este tipo es difícil, pero la melodía... es fácil de aprender. (*apud* Wray 115)⁷⁸

Como puede deducirse, los niños con un estilo expresivo son los que más utilizan un LF o prefabricado:

mientras que los niños referenciales poseen un vocabulario principalmente compuesto por sustantivos y con muy poco dominio de la morfología, los niños expresivos producen con la misma facilidad, dentro de las secuencias formulaicas, sustantivos, verbos, adjetivos y pronombres, así como también afijos correctos.⁷⁹ (Wray 2002: 114)⁸⁰

Un buen método para identificar que estilo cognitivo tiene el niño es contabilizar las estructuras que utiliza: «la proporción de frases arraigadas (*frozen*) adquiridas por el niño en las primeras y segundas 50 palabras es un buen indicativo para determinar un estilo no referencial» (*apud* Wray 2002: 115)⁸¹. Además, Wray

⁷⁷ Referentiality requires individual word labels for objects ... Expressiveness, on the other hand, requires a command of longer strings with an associated function. (Wray 2002: 114)

⁷⁸ Social-control language ... is likely to be heard in clumps that are not easily broken up; for example: 'D'ya wanna go out?'; 'I dunno where it is'; 'Stop it'. Segmentation of such sequences is difficult but the tune ... is easy to learn. (*apud* Wray 115)

⁷⁹ while referential children possess a vocabulary of mostly nouns, and very little command of morphology, expressive children produce with equal ease, within formulaic sequences, nouns, verbs, adjectives and pronouns, as well as correct affixes (2002: Wray 114).

⁸⁰ Una vez más, hay que aclarar que este uso de LF no significa que el niño se encuentre en un punto más avanzado del aprendizaje, o que tenga un mayor conocimiento gramatical. El uso de morfemas gramaticales está relacionado con la adquisición de frases como un conjunto, y en consonancia con Wray, «[Esto] no significa un conocimiento real de la gramática» (2002: 114). [«[it] does not signify real knowledge of the grammar» (2002: 114)].

⁸¹ «the proportion of frozen phrases acquired by the child in the first and second 50 words is a good candidate for a positive defining feature of non-referential style» (*apud* Wray 2002: 115).

expone que tener un determinado estilo cognitivo puede tener efectos a largo plazo una vez alcanzada la adultez. El sistema interno de reglas lingüísticas o la actitud hacia el aprendizaje de una segunda lengua pueden verse afectados por los estilos cognitivos adoptados en la infancia (Wray 2002: 114).

6.2. Aprendizaje en adultos (L2)

Uno de los aspectos más difíciles para un hablante no nativo es aprender a identificar los enunciados prototípicos que utilizan los hablantes y diferenciarlos de aquellos que *podrían* llegar a utilizar. Así, por ejemplo, no es lo mismo decir *lo tengo en la punta de la lengua* que *lo tengo al final de la lengua*, *metí la pata* que *introduje la pata*, o *en primer lugar* que *en primer sitio*. Aunque las segundas opciones son gramaticalmente posibles, no son adecuadas desde la perspectiva de un hablante nativo, e incluso un uso incorrecto puede provocar graves problemas comunicativos:

Errar al usar una expresión nativa puede producir una impresión de brusquedad, falta de respeto o arrogancia. (Wray 2002: 143)⁸²

las secuencias formulaicas utilizadas por los hablantes nativos no son fáciles de identificar y dominar para los alumnos, y su ausencia contribuye en gran medida a que los alumnos no suenen idiomáticos. (*ibid.* 176)⁸³

Debido a su gran importancia en la comunicación, Wray afirma que aprender pequeños conjuntos de secuencias formulaicas puede servir como una especie de puente comunicativo entre los hablantes nativos y los no nativos (2002: 143-4). Lo sostenido por Wray no difiere de lo que en los años noventa planteó Lewis con su enfoque léxico:

El enfoque léxico se funda, desde una perspectiva psicolingüística, en el reconocimiento de que, como en el caso de los hablantes nativos, una parte muy importante de la adquisición de una L2 consiste en la capacidad de comprender y producir frases léxicas como un todo sin analizar (*chunks*). La exhibición a estos segmentos léxicos permite a los alumnos asimilar de modo implícito patrones lingüísticos que

⁸² «Failing to use a nativelike expression can create an impression of brusqueness, disrespect or arrogance» (Wray 2002: 143)

⁸³ the formulaic sequences used by native speakers are not easy for learners to identify and master, and that their absence greatly contributes to learners not sounding idiomatic. (Wray 2002: 176)

se han considerado tradicionalmente parte de la gramática. La enseñanza se centra en expresiones bastante fijas que ocurren con frecuencia en el lenguaje hablado. (Marina Paz 427)

En primer lugar, las secuencias de LF sirven como una estrategia comunicativa en las primeras etapas de aprendizaje de una L2. En «Idiomaticity as an indicator of second language proficiency» (1989), Yorio investiga el uso de lo idiomático (fórmulas, colocaciones, etc.) en estudiantes de inglés como lengua extranjera. El estudio expone cómo los estudiantes con peores resultados tienen fallos gramaticales y léxicos en la producción de LF. Yorio interpreta que estas deficiencias formales en el control del lenguaje prefabricado son un indicativo para determinar el nivel de competencia comunicativa de los hablantes no nativos:

... parece que, aunque la fluidez es posible sin precisión gramatical, la idiomática no lo es. La idiomática se convierte así en un excelente indicador del dominio del sistema bilingüe y, como tal, merece ser estudiado y comprendido más a fondo. (1989: 68)⁸⁴

Yorio concluye que, al igual que hemos visto con los niños, el uso de expresiones idiomáticas no representa el conocimiento gramatical. De hecho, no propicia el desarrollo de este conocimiento:

La adquisición de formas convencionalizadas de lenguaje por parte de adultos que aprenden un segundo idioma parece ser una tarea larga y ardua... al principio, no parecen hacer un uso extendido de lenguaje prefabricado, formulaico, y cuando lo hacen, no parecen ser capaces de utilizarlo para incentivar su competencia gramatical. No es de extrañar que el lenguaje convencionalizado sea difícil de adquirir y utilizar si creemos que la fijación contribuye a la dificultad. (1989: 68)⁸⁵

Podemos asumir así que, cuando se adquiere, el lenguaje prefabricado es utilizado como un salvoconducto comunicativo en una L2, resultando más útil en lo práctico que en lo teórico. En este sentido, Wray afirma que, junto con

⁸⁴ ... it seems that although fluency is possible without grammatical accuracy, idiomaticity is not. Idiomaticity then becomes an excellent indicator of bilingual system proficiency and, as such, it deserves to be further studied and understood (1989: 68)

⁸⁵ The acquisition of conventionalized language forms by adults learning second language appears to be a long and arduous task ... they do not appear to make extensive early use of prefabricated, formulaic language, and when they do, they do not appear to be able to use it to further their grammatical development. It is not surprising that conventionalized language should be difficult to acquire and use if we believe that markedness contributes to difficulty. (1989: 68)

los objetivos personales del aprendiente, la motivación por integrarse con los hablantes de la lengua meta es un factor que promueve el uso de un LF: «parece haber un vínculo entre el uso de secuencias formulaicas y la necesidad y el deseo de interactuar. Estos dos aspectos contribuyen al logro general de la competencia comunicativa» (2002: 175)⁸⁶. El estudio de Yorio apoya esta idea. El autor muestra cómo los individuos con una menor distancia social y psicológica respecto a la cultura meta producen un lenguaje mucho más formulaico que aquellos que sufren un mayor distanciamiento social (1989: 60). La inmersión, como es ampliamente reconocida, desempeña un papel fundamental, tanto en el aprendizaje de la lengua en conjunto, como en el aprendizaje de lo idiomático y prefabricado en particular. También, el uso de un LF aporta un punto de fluidez muy importante para un hablante no nativo. Al interpretar la lengua como una sucesión de estructuras encadenadas, se produce una menor saturación en el procesamiento lingüístico, ya que las estructuras se encuentran almacenadas en la memoria y se accede rápidamente a ellas. Esta idea ya fue apuntada por Lewis en el enfoque léxico:

Lo que nos ayuda a hablar con fluidez es nuestra habilidad para usar grupos sintagmáticos [(lexical phrases)]. Este discurso prefabricado nos permite recuperar la información de forma más eficiente y también permite a los hablantes (y a los oyentes) dirigir su atención hacia la estructura general del discurso, en lugar de centrarse en las palabras individuales a medida que se van produciendo. Todo esto encaja perfectamente con los resultados obtenidos en la investigación computacional y en la investigación sobre la adquisición del lenguaje. (*apud* 1993: 19)⁸⁷

El aprendiente por lo tanto logra una mayor fluidez oral, menos titubeos y realiza pausas más coherentes entre enunciados. Como Wray apunta, «los fragmentos memorizados pueden no solo ser un síntoma de fluidez, sino también la causa de ella» (2002: 190)⁸⁸. Proporciona, además, un aumento en la confianza en la destreza oral, aunque es necesario destacar que, si se convierte en una tendencia habitual, puede estancar el proceso de aprendizaje.

⁸⁶ «there seems to be a link between the use of formulaic sequences and a need and desire to interact. These two together contributing to the overall achieve of communicative competence» (2002: 175).

⁸⁷ It is our ability to use lexical phrases that helps us speak with fluency. This prefabricated speech has both the advantages of more efficient retrieval and of permitting speakers (and hearers) to direct their attention to the larger structure of the discourse, rather than keeping it narrowly focussed on individual words as they are produced. All this fits very neatly with the results of computational and language acquisition research. (*apud* 1993: 19)

⁸⁸ «memorized chunks may not only be symptomatic of fluency but also the cause of it» (2002: 190).

En segundo lugar, el LF no solo sirve como estrategia comunicativa, sino también como estrategia didáctica. Los aprendientes pueden utilizar técnicas de aprendizaje que incluyan el uso de un LF: imitar los enunciados emitidos por el profesor, realizar prácticas mecánicas con construcciones, memorizar textos enteros o hacer acopio de secuencias usadas por los nativos (Wray 2002: 176). Asimismo, los manuales también pueden fundamentarse en secuencias de LF para la enseñanza de una lengua extranjera. Actualmente, es habitual encontrar manuales con un enfoque comunicativo que implementen la enseñanza de secuencias desde un primer momento. Esto ocurre especialmente en los niveles bajos e intermedios, donde es recurrente encontrar fórmulas de interacción básicas como *¿cómo te llamas?*, *¿de dónde eres?*, *¿a qué te dedicas?*, etc. Otro conocido ejemplo son los famosos métodos que aseguran alcanzar un alto grado de competencia mediante la memorización de un lenguaje prefabricado, dejando de lado la gramática —véase el caso del método Maurer, comúnmente conocido como «el inglés con mil palabras»—. Esta proliferación de técnicas y métodos basados en la memorización de estructuras demuestra la importancia que ha cobrado el lenguaje prefabricado y el fenómeno de lo idiomático bajo el contexto del enfoque comunicativo en la enseñanza. Como se ha mostrado anteriormente en el estudio de Olímpio de Oliveira Silva, las fórmulas tienen un alto valor funcional y así ha sido recogido en distintos métodos de enseñanza basados en esta premisa.

No obstante, resulta necesario remarcar que el lenguaje prefabricado no puede ser el pilar en la enseñanza de una lengua extranjera. Su valor funcional no debe ser puesto en duda, especialmente durante las primeras etapas de aprendizaje, pero es insuficiente para cubrir todas las necesidades comunicativas de un hablante. De acuerdo con Wray,

el uso temprano de secuencias formulaicas con necesidades comunicativas, [es un período] de corta duración, ya que las ambiciones comunicativas de los estudiantes pronto superan las limitaciones de las formas fijas. (2002: 181)⁸⁹

Si su capacidad solo se limita al uso de un lenguaje prefabricado, el aprendiente puede incluso llegar a utilizarlo como mecanismo para evitar ciertas situaciones que requieran una competencia lingüística alta:

⁸⁹ the early use of formulaic sequences with communicative needs, [is a] short lived [period], as the communicative ambitions of the learners soon outstrip the limitations of fixed forms (2002: 181).

los estudiantes de una L2 pueden fosilizar cualquier cadena de palabras útil una vez que se ha comprobado que funciona, incluso si no es nativa, o incluso pueden manipular específicamente su comportamiento para evitar ciertas situaciones que serían lingüísticamente difíciles. (Wray 2002: 148)⁹⁰

Asimismo, un excesivo uso de lenguaje prefabricado lleva a producir un habla inauténtico a ojos de un hablante nativo. En numerosas ocasiones, este uso excesivo es debido a las interferencias lingüísticas provocadas por la L1, aunque Wray afirma que estas interferencias están más o menos bajo el control del aprendiente: «la interferencia de la L1 está al menos parcialmente bajo el control del alumno, y no es solo una «fuga» inevitable causada por procesos psicolingüísticos» (2002: 183)⁹¹. Aun así, un alto grado de semejanza entre la L1 y L2 ayuda a cometer un menor número errores y a tener más confianza a la hora de utilizar fórmulas (Wray 2002: 183). Los aprendientes, además, pueden incurrir en el problema de reproducir secuencias en un contexto incorrecto, lo cual evidencia el importante factor pragmático e idiosincrásico de lo idiomático y prefabricado.

Finalmente, las investigaciones sugieren que la mejor opción es mantener un equilibrio entre la creatividad y lo prefabricado. Por un lado, la creatividad que se consigue a partir de un conocimiento gramatical proporciona un lenguaje amplio, dinámico y variado. Por otro, el uso correcto de secuencias prefabricadas aporta naturalidad y sencillez:

Sin un sistema basado en reglas, el repertorio del lenguaje sería limitado, cliché y, aunque adecuado para ciertos tipos de interacción, carecería de imaginación y novedad. Por el contrario, con solo un sistema basado en reglas, el lenguaje sonaría pedante, poco idiomático y vulgar. (*apud* Wray 2002: 183)⁹²

El problema para un hablante no nativo es, claramente, alcanzar esta armonía. La tónica habitual es que los hablantes nativos produzcan menos construcciones prefabricadas, pero más variadas, mientras que los hablantes no nativos tiendan

⁹⁰ L2 learners may cut their losses by fossilizing any useful string once it has been found to work, even if it is not natively like, or may even specifically manipulate their behaviour to avoid certain situations which would be linguistically difficult. (Wray 2002: 148).

⁹¹ «interference from the L1 is at least partially under control of the learner, not just unavoidable 'leakage' caused by psycholinguistic processes» (2002: 183).

⁹² «Without the rule-based system, language would be limited in repertoire, clichéd, and, whilst suitable for certain types of interaction, lacking imagination and novelty. In contrast, with only a rule-based system, language would sound pedantic, unidiomatic and pedestrian» (*apud* Wray 2002: 183).

a estancarse en el uso de las que ya tienen memorizadas y/o a fosilizar secuencias incorrectas. Este problema se debe a la naturaleza de una gran parte del lenguaje prefabricado: fórmulas rutinarias, proverbios y refranes, colocaciones, locuciones, etc. son estructuras de carácter espontáneo e idiosincrásico, por lo que son realmente difíciles de adquirir en el aula: «Al igual que sucede con los modismos y las rutinas, las colocaciones solo se pueden aprender si nos encontramos con ellas» (Wray 2002: 183)⁹³. Además, en numerosas ocasiones un hablante nativo procurará evitar el uso de construcciones ante un extranjero, por lo que el estancamiento en la adquisición de lo idiomático se acentúa. La exposición, por lo tanto, a un input variado resulta fundamental para aprender estas estructuras, y debe ser acompañado de una enseñanza de la gramática para proporcionar una mayor libertad a la hora de comunicarse. De acuerdo con Wray, únicamente de esta manera es posible llegar al proceso de segmentación, esto es, el punto del aprendizaje en el cual el aprendiente comienza a crear de manera autónoma nuevas secuencias —y, además, adecuadas— a partir de las secuencias previamente adquiridas y a un conocimiento gramatical. Wray observa lo siguiente en un estudio sobre el uso de fórmulas en aprendientes de una L2:

En las primeras etapas, las secuencias formulaicas ayudaron con la comunicación, pero luego sus necesidades comunicativas sobrepasaron lo que las secuencias de fórmulas podían ofrecer. (por ejemplo, la necesidad de hablar sobre terceras personas cuando las secuencias formulaicas estaban en segunda o primera persona). En respuesta, ellos/as «gradualmente ‘desempaquetaron’ sus primeros *chunks* y también usaron partes de ellos de manera productiva en la producción de nuevas expresiones». Las secuencias formulaicas no se eliminaron, sino que se conservaron y modificaron. Al principio, simplemente etiquetaron al nuevo referente con la misma fórmula sin cambios, pero luego descompusieron esta fórmula gradualmente. (2002: 187)⁹⁴

⁹³ «As with idioms and routines, collocations can only be learned if they are encountered» (Wray 2002: 183).

⁹⁴ In the early stages, the formulaic sequences helped with communication, but later their communicative needs overtook what the formulaic sequences could offer. (e.g., needing to talk about third persons when the formulaic sequences were second or first person). In response, they “gradually ‘unpacked’ their early chunks, and also used parts of them productively in the generation of new utterances”. The formulaic sequences weren’t dropped but retained and modified. At first, they simply tagged the new referent onto the unchanged formula, but then gradually broke them down. (2002: 187)

De esta manera, y al igual que la gramática, el LF tiene un papel fundamental, tanto en la competencia comunicativa, como en el proceso de aprendizaje en sí mismo:

Sigue siendo muy plausible que las secuencias formulaicas apoyen el proceso de adquisición, ya sea simplemente manteniendo en el alumno la sensación de poder decir algo, incluso cuando solo hay una pequeña base de datos a la que recurrir, o proporcionando una gran cantidad de datos almacenados sobre la lengua meta para su posterior análisis. (Wray 2002: 188)⁹⁵

⁹⁵ it remains highly plausible that formulaic sequences are supporting the acquisition process, whether this be simply by maintaining in the learner a sense of being able to say something, even when there is only a small database to draw on, or by providing a wealth of stored nativelike data for later analysis. (Wray 2002: 188)

7. Nuevas líneas de investigación: la lingüística computacional

Como ha sido expuesto en esta monografía, el estudio de lo idiomático y lo prefabricado ha cobrado una gran importancia en las últimas tres décadas. En este sentido, un campo de investigación cada vez más presente es la lingüística computacional. La lingüística computacional puede definirse como un campo interdisciplinar que investiga el procesamiento automático del lenguaje natural, en el que intervienen lingüistas, informáticos, psicolingüistas y expertos en lógica e inteligencia artificial⁹⁶. De acuerdo con Llisterri y Moure,

Bajo la denominación de *Lingüística computacional* es posible agrupar un conjunto relativamente heterogéneo de teorías, métodos, herramientas, aplicaciones y productos que tienen en común la consideración de la lengua como un objeto susceptible de ser tratado mediante procedimientos informáticos. (1996: 147)

El objetivo de esta línea de investigación es detectar formalismos y patrones en la lengua y adaptarlos a un lenguaje informático. En otras palabras, la lingüística computacional explora las interacciones entre las computadoras y el lenguaje humano, relacionando la lengua formal (informática) y la lengua natural (humana). Los formalismos y patrones lingüísticos que identifica la lingüística computacional pueden estar basados en reglas —es decir, en la gramática—, o basados en la estadística —es decir, en la probabilidad de que unas palabras aparezcan juntas en un conjunto—. Bajo este último criterio, la lingüística computacional, al igual que algunos procedimientos de identificación de LF, hace

⁹⁶ Son notablemente importantes para la lingüística computacional los trabajos realizados por L. T. F. Gamut, seudónimo utilizado por los expertos en lógica, lingüística y filosofía van Benthem, Groenendijk, de Jongh, Stokhof y Verkuyl. Las iniciales L. T. F. representan la lógica, el lenguaje y la filosofía (en alemán *logica*, *taal* y *filosofie* respectivamente), mientras que Gamut es el acrónimo de las universidades de Groninga (G), Amsterdam (am) y Utrecht (ut).

un uso aplicado de la lingüística de corpus. La lingüística computacional por lo tanto, también se preocupa por identificar aquellos aspectos del lenguaje que son gramaticalmente esquivos, pero que igualmente son una parte fundamental del lenguaje humano. A este respecto, Llisterri y Moure afirman:

El terreno de la *Lingüística computacional* puede cifrarse en la elaboración de productos derivados del enfoque teórico pero, al tiempo, esa elaboración debe revertir en la calidad de nuestros conocimientos. La *Lingüística computacional* parece un terreno idóneo para la verificación de teorías lingüísticas ... porque está poniendo en evidencia fenómenos de las lenguas escasamente conocidos y, a medida que pretenda diseñar gramáticas para autómatas irá desvelando nuevas cuestiones ... la lingüística de corpus [como rama de la lingüística que se aplica en la lingüística computacional], al proporcionar ingentes cantidades de datos reales sobre las lenguas, ha mostrado que las gramáticas son bastante menos regulares de lo que las teorías suelen considerar; que la aparente armonía de los sistemas lingüísticos, sustentada sobre un haz de oposiciones bien delimitado, se resquebraja ante la reiterada aparición de fenómenos periféricos, que no obedecen a reglas simples. (1996: 213-4)

En definitiva, la lingüística computacional tiene un importante papel en, por un lado, comprender mejor el funcionamiento del lenguaje humano y, por otro, en la difícil labor de integrar la tecnología en la vida cotidiana de las personas de la forma más comunicativa posible. En este último caso, destaca la aplicación de la lingüística computacional en el desarrollo de asistentes en sistemas informáticos, de correctores en procesadores de texto, de programas de reconocimiento del habla o de traductores automáticos:

Adoptando una perspectiva muy amplia, entrarán dentro de este ámbito [la lingüística computacional] desde las ayudas proporcionadas por los ordenadores a la investigación lingüística en disciplinas como la filología, la estilística, la estadística lingüística o la lexicografía hasta los sistemas capaces de llevar a cabo automáticamente la traducción de un texto o de ofrecer servicios telefónicos sin la intervención de un operador humano, pasando por los correctores que habitualmente se integran en los programas de tratamiento de texto o por los sistemas de recuperación automática de la información que hacen uso de datos lingüísticos. (Llisterri y Moure 1996: 147)

8. Conclusiones

Esta monografía ha pretendido dar a conocer la realidad del LF con un acercamiento general sobre la historia de su investigación, sus características principales, y su importante papel en lo relativo a la adquisición de una L1 y en el aprendizaje de una L2. Durante mucho tiempo lo idiomático y prefabricado en la lengua ha sido definido como un fenómeno periférico. Esta concepción fue ampliamente asumida durante la etapa en la que las teorías estructuralistas y chomskianas conformaban el paradigma lingüístico y educativo.

Sin embargo, con el auge del enfoque comunicativo a partir de los años ochenta y los avances en distintas disciplinas como la psicolingüística, la lingüística cognitiva o la pragmalingüística se produjo un cambio respecto la concepción de lo prefabricado en el discurso, otorgando a este aspecto de la lengua un rol más importante. Muestra de ello son las teorías lingüísticas de carácter interdisciplinar que surgieron hacia finales del siglo XX. Tanto la gramática de construcciones como el LF son teorías lingüísticas que adoptan un enfoque de la lengua basado en el uso, y que, consecuentemente, nos hacen replantearnos la recurrente mentalidad que asocia únicamente la competencia en una L2 o LE con la competencia lingüística o gramatical.

A pesar de las dificultades que entraña la investigación del LF—aspectos como su identificación o su taxonomía resultan bastante complejos—, las secuencias de LF son conjuntos léxicos —*chunks*— de vital importancia en la comunicación. Algunas de sus funcionalidades más importantes son la reducción en la carga de procesamiento, el aumento de la fluidez y su valor como estrategia comunicativa en una L2 o incluso como recurso didáctico. Además, desde un punto de vista etnolingüístico las fórmulas tienen un carácter ritual cuya función culminante es armonizar la vida del individuo en sociedad, promoviendo la dimensión del aprendiente como agente social.

En definitiva, el LF se relaciona con la competencia comunicativa que tan presente está en la enseñanza de lenguas extranjeras desde los años ochenta. Como Martí Sánchez apunta, «la competencia comunicativa es una realidad heterogénea donde convergen lo específicamente lingüístico, lo cultural y los aspectos universales de la cognición humana, por eso no presenta el mismo grado de definición que la competencia gramatical» (9). Las informaciones e ideas que en esta monografía se han presentado apuntan a relacionar lo formulaico con esa idea de heterogeneidad. En el LF se concatena lo estrictamente lingüístico, lo idiosincrásico o social y lo innato o filogenético al ser humano.

Referencias bibliográficas

Libros

- GOLDBERG, Adele E. *Constructions: A Construction Grammar Approach to Argument Structure*. The University of Chicago Press, 1995.
- LEWIS, Michael. *The Lexical Approach. The State of ELT and a Way Forward*. Language Teaching Publications, 1993.
- MAKKAI, Adam. *Idiom Structure in English*. Mouton Publishers, The Hague, 1972.
- PETERS, Ann M. *The Units of Language Acquisition*. Cambridge University Press, 1983.
- SALVADOR, Vicent. *El discurs prefabricat. Estudis de fraseologia teòrica i aplicada*, editado por Adolf Piquer. Universitat Jaume, 2000.
- WOOD, David. *Fundamentals of Formulaic Language. An introduction*. Bloomsbury Publishing Plc, 2015.
- WRAY, Alison. *Formulaic Language and the Lexicon*. Cambridge University Press, 2002.

Artículos

- BUERKI, Andreas. «Formulaic sequences: a drop in the ocean of constructions or something more significant?». *European Journal of English Studies*. vol. 20, Issue 1, 2016, pp. 15-34.
- CALZADO, Araceli Roldán. «Las construcciones en la clase de español como lengua extranjera». *MarcoELE*. no. 21, Jul – Dec 2015.
- CHAFE, Wallace L. «Idiomatcity as an Anomaly in the Chomskyan Paradigm». *Foundations of Language*. vol. 4, no. 2, May 1968, pp. 109-127.
- CONTRERAS IZQUIERDO, Narciso M, y Ruth M^a Rey Arranz. «Hacia un modelo de análisis multidimensional de las secuencias formulaicas para su aplicación en ELE/EL2». *Revista Electrónica de Lingüística Aplicada*. vol. 17, no. 1, 2018, pp. 20-44.
- ERMAN, Britt y Beatrice Warren. «The idiom principle and the open choice principle». Walter de Gruyter, *Text 20 (1)*, 2000, pp. 29-62.

- FILLMORE, Charles J, et al. «Regularity and Idiomaticity in Grammatical Constructions: The Case of Let Alone». *Language*, vol. 64, no. 3, Sep 1988, pp. 501-538.
- GONZÁLEZ REY, M^a Isabel. «De la didáctica de la fraseología a la fraseodidáctica». *Paremia*. no. 21, 2012. pp. 67-84.
- HUDSON, Richard A. Crítica de *Idiom Structure in English*, de Adam Makkai. *Journal of Linguistics*, vol. 11, no. 1, Mar 1975, pp. 73-80.
- LLISTERRI, Joaquim, y Teresa Moure. «Lenguaje y nuevas tecnologías: el campo de la lingüística computacional». *Avances en lingüística aplicada*. 1996, pp. 147-228.
- MARINA PAZ, Gabriela. «El enfoque léxico: una perspectiva innovadora en la enseñanza de ELE». Centro Virtual Cervantes. Thomson Learning, Edinumen, 2007.
- MARTÍ SÁNCHEZ, Manuel. «La competencia fraseológica como objeto problemático de la fraseodidáctica». *Innovación en fraseodidáctica: tendencias, enfoques y perspectivas*, editado por Carlos A. Crida Álvarez, Peter Lang, 2019, pp. 11-23.
- OLÍMPIO DE OLIVEIRA SILVA, María Eugênia. «Aportaciones de la Etología Humana a los estudios lingüísticos: el caso de la Fraseología». *Pragmalingüística*. no. 23, Dec 2015, pp. 151-170.
- PAWLEY, Andrew. «Developments in the study of formulaic language since 1970: a personal view». *Phraseology and Culture in English*, editado por Paul Skandera, Mouton de Gruyter, 2007, pp. 3-45.
- WEINREICH, Uriel. «Problems in the Analysis of Idioms». *Substance and Structure of Language*, editado por Jaan Puhvel, University of California Press, 1969, pp. 23-81.
- WRAY, Alison, y Michael R. Perkins. «The functions of formulaic language: an integrated model». *Language and Communication*, vol. 20, no. 1, Jan 2000, pp. 1-28.
- YORIO, Carlos A. «Idiomaticity as an indicator of second language proficiency». *Bilingualism Across the Lifespan: Aspects of Acquisition, Maturity and Loss*, editado por Kenneth Hyltenstam, y Lorraine K. Obler. Cambridge University Press, 1989, pp. 55-69.

Capítulos de libros

- DIESEL, Holger. «Construction grammar and first language acquisition». *The Oxford Handbook of Construction Grammar*, editado por Thomas Hoffmann, y Graeme Trousdale. Oxford University Press, 2013.

Tesis

- HERNÁNDEZ MUÑOZ, Yaiza Irene. *Las construcciones francesas fundamentales: definición y aplicación de una nueva unidad fraseológica*. Tesis. Universidad Complutense de Madrid, 2019.
- MORENO TEVA, Inmaculada. *Las secuencias formulaicas en la adquisición de español L2*. Tesis. Universidad de Estocolmo, 2012. Disponible en: <https://hispanismo.cervantes.es/publicaciones/las-secuencias-formulaicas-adquisicion-espanol-l2>.

STUKOVA, Natalya. *The Role of Formulaic Language in the Creation of Grammar*. Tesis. University of New Mexico, 2011. Disponible en: https://digitalrepository.unm.edu/ling_etds/33.

Webs

Diccionario de términos clave de ELE. «Competencia gramatical». Centro Virtual Cervantes. Disponible en: https://cvc.cervantes.es/Ensenanza/biblioteca_ele/diccio_ele/default.htm.

—. «Enfoque comunicativo». Centro Virtual Cervantes.

—. «Cortesía». Centro Virtual Cervantes.

Plan Curricular Del Instituto Cervantes. Alcalá de Henares, Madrid: Instituto Cervantes, Biblioteca Nueva, 2006. Disponible en: https://cvc.cervantes.es/ensenanza/biblioteca_ele/plan_curricular/.

Anexos

Anexo 1

Términos utilizados en la investigación del LF para describir las secuencias formulaicas (*apud* Wray y R. Perkins 2000: 3):

amalgams	gambits	preassembled speech
automatic	gestalt	prefabricated routines and patterns
chunks	holistic	patterns
clichés	holophrases	ready-made expressions
co-ordinate constructions	idiomatic	ready-made utterances
collocations	idioms	rote
composites	irregular	routine formulae
conventionalized forms	lexical(ised) phrases	schemata
FEIs	lexicalised sentence stems	semi-preconstructed phrases
fixed expressions	multiword units	that constitute single choices
formulaic language	non-compositional	sentence builders
formulaic speech	non-computational	stable and familiar expressions
formulas/formulae	non-productive	with specialized subsenses
fossilized forms	petrification	synthetic
frozen phrases	praxons	unalused chunks of speech

E-**le**ando

ELE en Red.

Serie de monografías y materiales para la enseñanza de ELE

E- de enseñanza y excelencia, esencia de nuestra identidad; ELE de Español como Lengua Extranjera, nuestra razón de ser. *E-**le**ando* es una publicación periódica y digital que, aprovechando las ventajas de la red, pone a disposición de todo aquel que esté interesado en el Español como Lengua Extranjera (ELE), en cualquier punto del planeta, una serie de monografías y materiales didácticos realizados por expertos en las diversas materias que integran este apasionante ámbito de la lingüística aplicada.

*E-**le**ando. ELE en Red* es editada por el [Máster en Enseñanza de Español como Lengua Extranjera](#), estudio propio de la Universidad de Alcalá, y por el [Máster Universitario en Formación de Profesores de Español](#); con ellos, la Universidad de Alcalá atesora más de 20 años formando especialistas en ELE. Nace en 2017, bajo la dirección de Ana M. Cestero Mancera, catedrática de Lingüística General de la Universidad de Alcalá, y cuenta con un comité de expertos integrado por profesionales de gran prestigio.

El lenguaje formulaico y su importancia en la adquisición/aprendizaje de segundas lenguas

Lo idiomático y prefabricado ha despertado el interés de diversas disciplinas y corrientes a lo largo de las últimas tres décadas, especialmente a partir del asentamiento del enfoque comunicativo como modelo didáctico. El lenguaje formulaico es una de las teorías más recientes que adopta un enfoque de la lengua basado en el uso, asemejándose así a lo propuesto por la más conocida gramática de construcciones. Al igual que la gramática de construcciones, la teoría del lenguaje formulaico tiene un carácter interdisciplinar, y, a pesar de su escasa presencia en la lingüística española, sus contribuciones arrojan unas ideas que pueden resultar muy útiles para seguir avanzando en el conocimiento sobre el funcionamiento de la lengua, así como en su enseñanza.

