

Aproximación a la fraseología de los inmigrantes (comparación de los materiales del corpus ISPIE y del corpus PRESEEA-Granada)

Marcin Sosinski

El artículo presenta los resultados del análisis cuantitativo de las unidades fraseológicas empleadas por los inmigrantes marroquíes residentes en Granada y cuyas entrevistas forman parte del corpus ISPIE. Se examinan cuatro aspectos: la frecuencia de aparición de fraseologismos en general, la frecuencia de aparición de unidades diferentes, la frecuencia de aparición de tipos de expresiones pluriverbales y la aparición de unidades fraseológicas concretas. Los resultados se comparan con un análisis similar efectuado sobre una entrevista de un hablante nativo que forma parte del corpus PRESEEA-Granada. Como resultado, se pueden observar diferencias entre el hablante español y los informantes extranjeros, y, por otro lado, entre los hablantes marroquíes.

An approach to the immigrants' phraseology (comparison of ISPIE corpus and PRESEEA-Granada corpus). The paper presents the result of a quantitative analysis of the phraseological units used by the Moroccan immigrants from the Spanish city of Granada and whose interviews are part of the ISPIE corpus. Four aspects are examined: the general frequency of appearance, the frequency of appearance of different units, the frequency of appearance of different types of units and the frequency of specific units. The results are compared with a similar analysis which has been carried out for one interview of a Spanish native speaker which is part of the corpus PRESEEA-Granada. Differences are observed between the native speaker and non native speakers, on one hand, and within the group of the non native speakers, on the other.

Palabras claves: fraseología, España, inmigración.

1. Introducción

Este trabajo tiene por objetivo analizar la fraseología, en concreto las locuciones y las fórmulas discursivas, empleada por los inmigrantes y compararla con las unidades pluriverbales empleadas por hablantes nativos del español.

Las grabaciones analizadas son fruto de dos proyectos de investigación en los que he tenido el placer de participar: el proyecto *Integración sociolingüística de la población inmigrante en España* (ISPIE) y el proyecto *Estudio Sociolingüístico del Español de Granada* (ESCEGRA, ref.: HUM2007-65602-C07-04/FILO) subvencionado por el Ministerio de Educación y Ciencia.

El proyecto granadino se inscribe en el proyecto panhispánico PRE-SEEA (*Proyecto para el Estudio Sociolingüístico del Español de España y América*), persigue sus mismos objetivos e incorpora su misma metodología con el fin de crear un corpus lingüístico de español hablado en Granada, obtenido a partir de muestras orales representativas. El corpus ya está terminado, editado (Moya Corral 2007, 2008, 2009) y puede consultarse en la página web del proyecto: <http://www.ugr.es/local/escegra>. En estos momentos nos estamos dedicando a su estudio, en función de los intereses científicos de los integrantes del grupo. De esta manera, la presente comunicación es una muestra de la investigación fraseológica que estamos llevando a cabo en Granada.

Por otro lado, mi vinculación con el proyecto *Integración sociolingüística de la población inmigrante en España* fue la de un investigador de campo encargado de grabar a informantes marroquíes residentes en Granada. Sin embargo, no me he quedado ajeno a los objetivos científicos del proyecto, entre los que figura el siguiente punto: “Realizar un análisis contrastivo de los repertorios lingüísticos y comunicativos de inmigrantes y receptores, con el fin de identificar los aspectos que ofrecen dificultades para la interacción comunicativa”.

En este caso, el trabajo aquí presentado puede considerarse como un ejemplo de un estudio contrastivo, centrado, obviamente, en la fraseología.

2. Aspectos metodológicos

Las posibilidades de estudio que ofrecen las unidades fraseológicas son muy amplias, no obstante, en esta ocasión, me he decidido por un estudio de corte cuantitativo, en el cual voy a examinar cuatro aspectos:

1. la frecuencia de aparición de unidades fraseológicas distintas;

2. la frecuencia de aparición de unidades fraseológicas consideradas en su totalidad;
3. la frecuencia de aparición de distintos tipos de unidades fraseológicas;
4. la frecuencia de aparición de unidades fraseológicas concretas.

Se efectuará una comparación entre los hablantes no nativos y el hablante nativo seleccionado como referente. Por otra parte, también se realizará la misma operación considerando solo a los inmigrantes, para analizar las diferencias pertinentes, en función de su competencia lingüística.

A este respecto, la impresión intuitiva, convertida en hipótesis, nos dice que:

1. El número de unidades distintas crecerá condicionado por el nivel del dominio del español: la más alta debería observarse en el hablante nativo y la más baja en el discurso de los inmigrantes de nivel inicial.
2. Paralelamente, la frecuencia de aparición de unidades fraseológicas crecerá condicionada por el nivel del dominio de la lengua: una vez más, la más alta debería observarse en el hablante nativo y la más baja en el discurso de los inmigrantes de nivel inicial.
3. El examen de la entrevista del hablante nativo nos revelará qué tipos de unidades fraseológicas aparecen con más frecuencia: suponemos que el estudio de los hablantes con el nivel más alto del dominio del español debe ofrecer datos parecidos.
4. El análisis de las unidades más frecuentes: al tratarse de un estudio con un número reducido de hablantes, creemos que esta parte de la investigación solo nos aportará información sobre los idiolectos de los informantes.

Para realizar el trabajo, hemos analizado las cuatro grabaciones realizadas en Granada, todas de informantes marroquíes con una estancia máxima de cinco años en España:

Mo1: hombre de unos 25 años, conocimiento del español: inicial
 Mo2: hombre de unos 25 años, conocimiento del español: inicial
 Mo3: hombre de unos 25 años, conocimiento del español: intermedio
 Mo4: mujer de unos 55 años, conocimiento del español: intermedio alto

Las grabaciones fueron transcritas y revisadas con la intención de analizar su componente fraseológico. La entrevista del sujeto 3, más

larga que las demás, fue recortada, descartándose la segunda parte. No obstante, aún se percibían diferencias entre las entrevistas que, a pesar de durar alrededor de una hora, contenían números de palabras diferentes, desde 5000 hasta casi 8000 vocablos.

Es una diferencia importante, capaz de distorsionar los resultados ya que los hablantes que más unidades léxicas producen también tienen más probabilidad de introducir en su discurso una unidad pluriverbal. Por ello, he creído conveniente estudiar no los números absolutos sino las frecuencias.

Por otra parte, como referencia, he escogido a uno de los informantes del proyecto de Granada, como muestra de una investigación más amplia que tiene como objetivo analizar las unidades fraseológicas en el corpus de los hablantes nativos.

Por último, me he fijado en las locuciones y en las fórmulas rutinarias que integran el grupo de las fórmulas discursivas, obteniendo, como resultado, un corpus de fraseologismos que fueron ordenados y reducidos a su forma canónica siguiendo la propuesta de Montoro del Arco (2004, 2005) y nuestras propias ideas expresadas en otra ocasión (Sosinski 2008).

3. Teoría fraseológica

La investigación ha sido realizada desde el campo de la fraseología y, para llevarla a buen puerto, ha sido necesario elegir un marco teórico apropiado. En este sentido, son de sobra conocidos los trabajos de, por ejemplo, Casares (1992[1950]), Zuluaga (1980), Ruiz Gurillo (1997), Castillo Carballo (1997), García-Page (2008) o la simplificada clasificación de la Real Academia que fue presentada en su última *Gramática* (2009). Aquí hemos optado por basarnos en la segmentación clásica del universo fraseológico que propuso en su momento Corpas Pastor (1996) y que fue posteriormente parcialmente reformulada por investigadores como Montoro del Arco (2006, 2008).

Veamos sus principios básicos. En primer lugar, estudiamos las locuciones, definidas por la autora malagueña (Corpas Pastor 1996: 88) como

unidades fraseológicas del sistema de la lengua con los siguientes rasgos distintivos: fijación interna, unidad de significado y fijación externa pasemática. Estas unidades no constituyen enunciados completos, y, generalmente, funcionan como elementos oracionales.

En la elaboración de una taxonomía de las locuciones españolas, Corpas Pastor sigue el criterio de la función oracional que pueden desempeñar estas unidades. Para resolver el problema de la existencia de las locuciones plurifuncionales, aparte del criterio mencionado, la autora también se fija en el núcleo del sintagma correspondiente, esto es, complementa el criterio funcional con el criterio formal.

Las locuciones nominales desempeñan las mismas funciones oracionales que un sustantivo: *mosquita muerta, lágrimas de cocodrilo*.

Las locuciones adjetivas, lógicamente, cumplen las mismas funciones que los adjetivos: *corto de medios, fuerte como un toro*.

Las locuciones adverbiales, un grupo muy numeroso y complejo, incluyen sobre todo los sintagmas prepositivos *a bordo, a brazo partido* aunque no es extraño encontrar entre ellas sintagmas adverbiales *más tarde o más temprano*, nominales *gota a gota* y adjetivales *largo y tendido*.

En un trabajo posterior (Corpas Pastor 1998), la autora se reafirma en la posibilidad de que se pudiese hablar de un tipo de locuciones adjetivo-adverbiales, cosa que Castillo Carballo (1997), en su tesis, manifiesta abiertamente.

Las locuciones verbales son las que, como los verbos, forman predicados: *poner el dedo en la llaga*.

Las locuciones prepositivas, diferentes de las preposiciones agrupadas y que pueden funcionar como nexos en sintagmas prepositivos, tienen la estructura de un adverbio (o un sustantivo adverbializado) más una preposición o de un sustantivo (o dos coordinados) más una preposición y, en algunos casos, también precedido por otra preposición: *en torno a*.

Las locuciones conjuntivas tienen, según la autora, un *estatus vacilante*. La clasificación que se hace de estos elementos es de tipo tradicional, distinguiendo entre las locuciones conjuntivas coordinantes y subordinantes, sin hacer una separación tajante entre los dos tipos: *tan pronto como*.

El último grupo de las locuciones está compuesto por las locuciones clausales que, a pesar de presentar la estructura de sujeto y predicado, no pueden funcionar como enunciados autónomos, ya que uno de sus actantes debe actualizarse en el discurso: *hacérsele a alguien la boca agua, no llegarle a alguien la camisa al cuerpo*.

Por otro lado, hay que hacer mención de la aportación de Esteban Montoro del Arco (2006, 2008) quien distingue las locuciones prepositivas, conjuntivas y marcadoras y las engloba todas bajo el rótulo de locuciones particulares que se caracterizan por su carácter gramatical.

En segundo lugar, incluimos en nuestro estudio las fórmulas rutinarias que forman parte de un grupo más amplio, el de *enunciados fraseológicos*, y quedan definidas como (Corpas Pastor 1996: 170)

UFs del habla, con carácter de enunciado, las cuales se diferencian de las paremias por carecer de autonomía textual, ya que su aparición viene determinada, en mayor o menor medida, por situaciones comunicativas precisas.

Nos interesan solo las fórmulas discursivas que tienen la función de organizar el discurso y que, en palabras de Corpas Pastor (1996: 189-190), funcionan

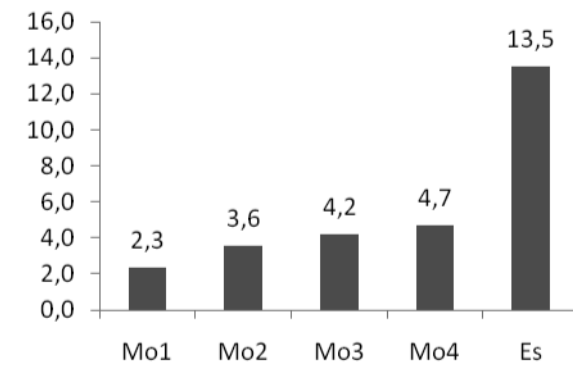
regulando la interacción, organizado y precisando lo que se dice, resaltando alguna parte, enlazando unos tópicos con otros, y permitiendo a los interlocutores tomar la palabra, mantener el turno u orientar el cambio de éste.

En lo que concierne a nuestro trabajo, una de las mayores dificultades consiste en el estatus vacilante de algunas unidades. Por ejemplo, los elementos del tipo *y ya está, y todo, y nada, y eso*, son clasificadas como locuciones o como fórmulas rutinarias, en función del grado de autonomía textual que les conceda el investigador. En este trabajo, las ubicamos en este último grupo, es decir, las analizaremos como fórmulas rutinarias.

Otro ejemplo de estas dificultades lo constituyen las locuciones marcadoras que tradicionalmente han sido clasificadas como elementos del grupo de las locuciones adverbiales u otras categorías. Así, Ruiz Gurillo, al incluir en su propuesta las locuciones marcadoras, elimina a la vez las locuciones conjuntivas. Es importante tener en cuenta estos hechos porque de las decisiones que hemos tomado al respecto dependen los resultados.

4. Los resultados y su análisis

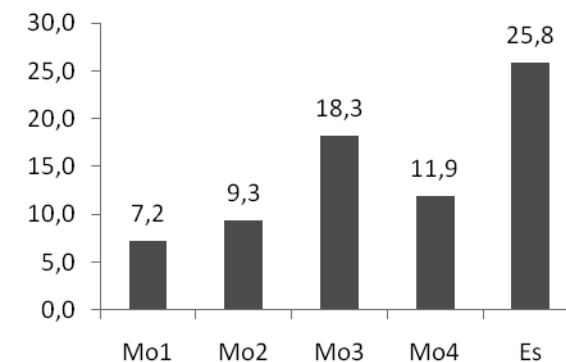
Se ha efectuado un recuento y un análisis de las locuciones y fórmulas rutinarias, obteniendo los siguientes resultados:



a) Frecuencia de aparición de unidades fraseológicas distintas (por 1000 palabras)

Se observa una clara diferencia entre el hablante español y los informantes marroquíes. El sujeto Mo4, que más unidades fraseológicas ha utilizado ni siquiera ha conseguido emplear la mitad del número de fraseologismos que detectamos en la entrevista del hablante nativo.

Por otra parte, también es significativa la diferencia entre los hablantes no nativos. Repárese en que, al ordenar los informantes en función de esta característica, hemos obtenido un orden que se corresponde exactamente con su nivel de competencia lingüística. Por ello, este tipo de análisis parece ser un instrumento adecuado para determinar el nivel del dominio del español por parte de un extranjero.



b) Frecuencia de aparición de unidades fraseológicas (por 1000 palabras)

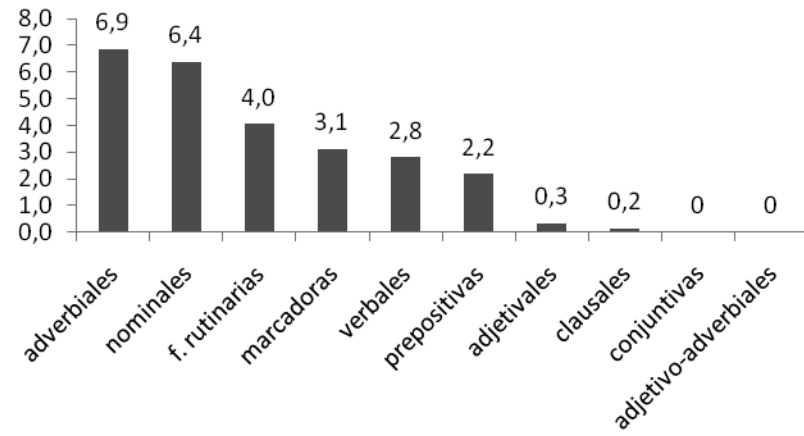
También en este punto se observa un contraste entre el hablante español y los informantes extranjeros, aunque las diferencias entre los inmigrantes del nivel inicial, medio y superior ya no son tan claras.

Ello se debe al uso repetitivo que algunos de ellos hacen de ciertas unidades que se convierten en muletillas que suplen sus carencias lingüísticas y discursivas. Por ejemplo, el informante Mo3, en cuyo discurso he encontrado 144 fraseologismos, ha empleado la locución *un poco* 60 veces y 43 veces otra locución adverbial, *a veces*. En el discurso del hablante Mo2, tres unidades, *de verdad*, *y ya está* y *no sé qué*, constituyen casi la mitad de los fraseologismos utilizados.

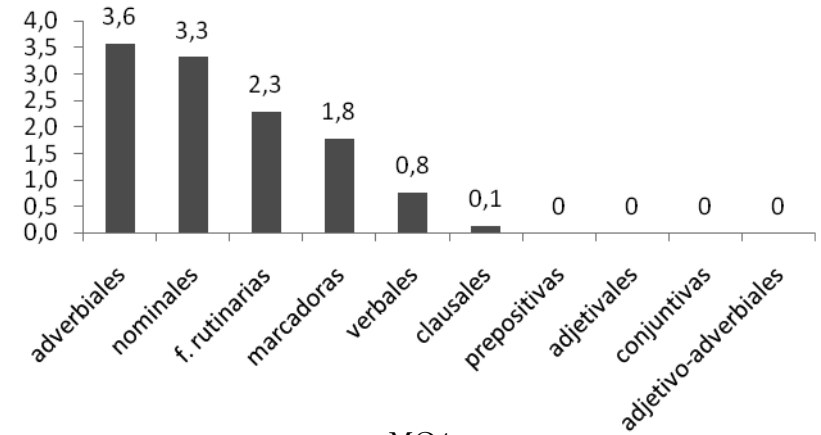
Por eso, antes de llegar a conclusiones precipitadas, conviene analizar si una alta frecuencia de aparición de expresiones pluriverbales, en vez de indicar un alto nivel del conocimiento de la lengua, no apuntan, precisamente, a lo contrario.

Esta faceta del análisis, si exceptuamos al informante Mo3, cuyo caso es un indicio de que es necesaria una reflexión metodológica más profunda, revela que la frecuencia de uso de fraseologismos crece a medida que los extranjeros van mejorando su nivel del español.

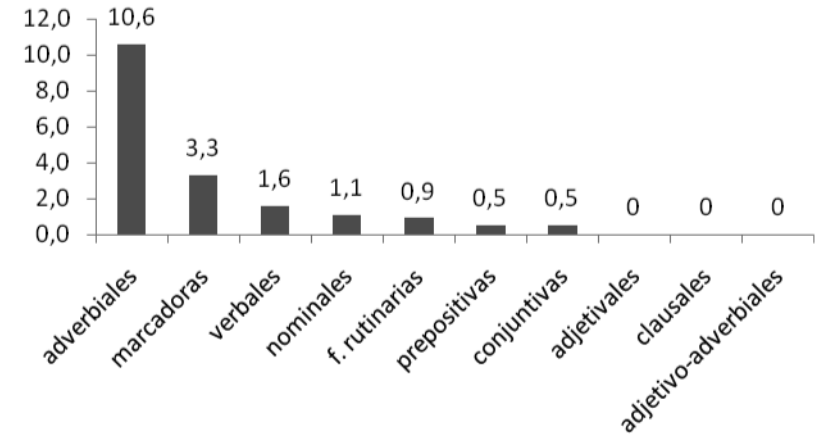
c) Frecuencia de aparición de distintos tipos de unidades fraseológicas (por 1000 palabras)



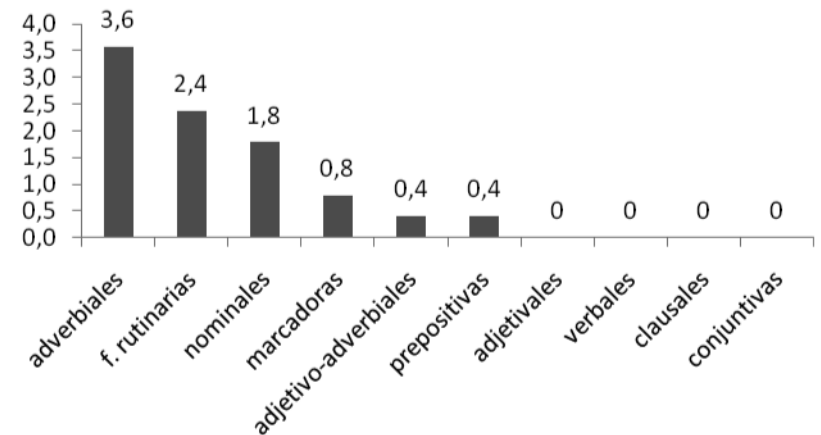
ES



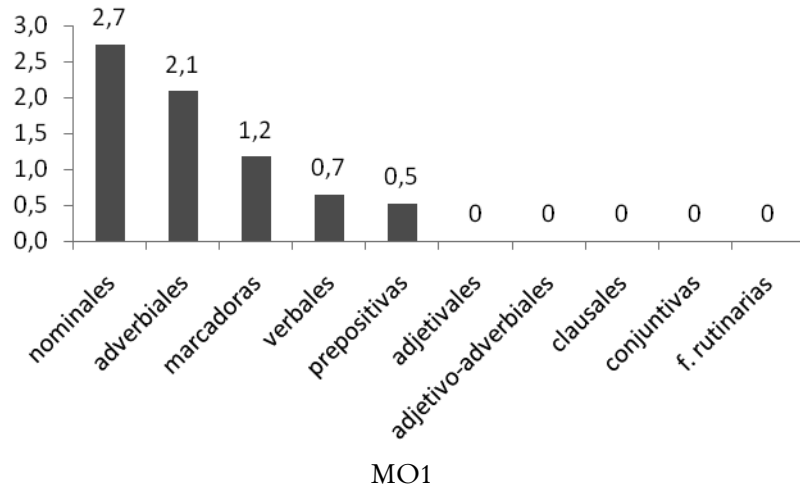
MO4



MO3



MO2



En este punto analizamos qué clases de unidades fraseológicas aparecen con más frecuencia en el discurso de los informantes. De nuevo, buscamos diferencias y semejanzas con respecto al hablante nativo. Así, se observa que en la entrevista del informante español aparece una gran variedad tipológica de unidades fraseológicas, con la distribución reflejada en la tabla. Destacan las locuciones adverbiales y las nominales, aunque hay que subrayar también el alto porcentaje de las marcadoras y de las fórmulas rutinarias, especialmente importantes para la estructuración de la comunicación y que, por lo tanto, influyen en la percepción de la pericia comunicativa.

Repárese que el recuento de los fraseologismos de la informante Mo4, la que más se aproxima a la competencia de un nativo, ofrece datos muy parecidos. De hecho, los cinco primeros tipos, aunque, lógicamente, con un índice de aparición más bajo, reproducen la misma secuencia: locuciones adverbiales, nominales, fórmulas rutinarias, locuciones marcadoras y verbales.

El análisis de la entrevista del hablante Mo3, con un nivel más básico del español, revela que, a pesar de que los cinco primeros grupos son los mismos, el orden, en función de la frecuencia de aparición, está sensiblemente alterado con respecto al hablante nativo. En el caso de los informantes Mo1 y Mo2, ni siquiera se mantiene la unidad de este conjunto de cinco tipos de fraseologismos más frecuentes.

Por último, cabe destacar que en todas las entrevistas las locuciones adverbiales ocupan un lugar muy relevante.

A la vez, no obstante, estos mismos datos nos indican la existencia de desajustes en el uso de algunas expresiones. Obsérvese la tabla correspondiente al informante Mo3 y el considerablemente alto índice de aparición de las locuciones adverbiales. Es un indicio que sugiere al investigador que el hablante abusa de este tipo de unidades, fenómeno adelantado en el punto anterior.

d) Unidades fraseológicas más utilizadas

Esta faceta de nuestra investigación, dado el reducido número de informantes analizados, tiene un valor limitado y debe relacionarse con aquellos estudios que, por ejemplo, se centran en la fraseología de una obra literaria o de un autor en concreto.

Aquí hemos optado por citar las tres unidades fraseológicas más frecuentes. Si varias unidades tienen el mismo índice de aparición, se citan todas.

Informante Es:

1. *al lado de* (particular: prepositiva),
2. *y eso* (fórmula rutinaria), *y ya está* (fórmula rutinaria), *fin de semana* (nominal), *por (la) parte de* (particular: prepositiva), *llamar la atención* (verbal)

Informante Mo4:

1. *todo el mundo* (nominal)
2. *y ya está* (fórmula rutinaria)
3. *en general* (particular: marcadora)

Informante Mo3:

1. *un poco* (adverbial)
2. *a veces* (adverbial)
3. *en principio* (particular: marcadora)

Informante Mo2:

1. *de verdad* (adverbial)
2. *y ya está* (fórmula rutinaria), *no sé qué* (nominal)

Informante Mo1:

1. *cada uno* (nominal)
2. *todo el mundo* (nominal)
3. *por la noche* (adverbial), *por eso* (particular: marcadora)

Repárese en que sólo en el caso del hablante nativo y de la informante Mo4 la mayoría de estas unidades está constituida por locuciones particulares y fórmulas rutinarias que, simplificando, sirven para enlazar elementos en el discurso.

6. Conclusiones

Esta breve comparación, aún reconociendo el carácter provisional de los resultados dada la escasa representatividad de la muestra, nos permite afirmar lo siguiente:

a) El hablante nativo ha utilizado en su discurso un número más alto de unidades fraseológicas que los informantes marroquíes. La misma diferencia se aprecia entre los extranjeros con distintos niveles del dominio de la lengua.

b) El hablante nativo ha recurrido a la fraseología con mucha más frecuencia que los informantes no nativos. La diferencia entre los hablantes marroquíes también es perceptible, pero los datos están desfigurados debido al uso recurrente de algunas expresiones.

c) Los cinco tipos de fraseologismos más utilizados, ordenados de mayor a menor frecuencia, son: locuciones adverbiales, nominales, fórmulas rutinarias, locuciones marcadoras y verbales. Este esquema queda alterado en los informantes con un nivel más bajo de competencia lingüística.

d) Entre las unidades fraseológicas más utilizadas, en el caso del hablante nativo y en el de la informante marroquí que mejor domina el español, destacan las locuciones particulares y las fórmulas rutinarias. Su alta frecuencia de aparición nos indica la importancia que tienen en la comunicación.

En resumen, el informante nativo ha utilizado muchas más unidades fraseológicas y lo ha hecho con más frecuencia que los informantes marroquíes. Se ha demostrado, pues, que hay una diferencia entre los hablantes no nativos y el hablante nativo.

Por otra parte, también es posible describir regularidades en función del grado de la competencia lingüística de los extranjeros. En este sentido, el criterio más fiable parece ser la frecuencia de aparición de unidades fraseológicas distintas. También puede ser útil el dato de la frecuencia del uso de expresiones pluriverbales, aunque es preciso idear un recurso metodológico especial que evite la contaminación de este dato por el uso excesivo de ciertas fórmulas.

Es evidente que este estudio tiene que ampliarse. Por un lado, en estos momentos estamos desarrollando una investigación parecida sobre el corpus PRESEEA-Granada. Por otra parte, los materiales del proyecto ISPIE también estarán disponibles en la red para su análisis. Por último, será interesante comparar estos resultados con los que arroje el examen de las entrevistas de alumnos extranjeros que aprenden el español en contextos formales.

Marcin Sosinski
 Universidad de Granada
 Departamento de Lengua Española
 Facultad de Filosofía y Letras
 Campus de Cartuja s/n
 18071 - Granada - España
 sosinski@ugr.es

Referencias bibliográficas

- Casares, Julio. [1950] 1992. *Introducción a la lexicografía moderna*. Madrid: Consejo Superior de Investigaciones Científicas.
- Castillo Carballo, María Auxiliadora. 1997. *Unidades pluriverbales en un corpus del español contemporáneo*. Tesis doctoral. Málaga: Servicio de Publicaciones de la Universidad.
- Corpas Pastor, Gloria. 1996. *Manual de fraseología española*. Madrid: Gredos.
- Corpas Pastor, Gloria. 1998. "Criterios generales de clasificación del universo fraseológico de las lenguas, con ejemplos en español y en inglés". En *Diccionarios, frases, palabras*, Manuel Alvar Ezquerro y Gloria Corpas Pastor (eds), 157-187. Málaga: Universidad de Málaga.
- García-Page Sánchez, Mario. 2008. *Introducción a la fraseología española: estudio de las locuciones*. Barcelona: Anthropos.
- Montoro del Arco, Esteban T. 2004. "La variación fraseológica y el diccionario". En *De Lexicografía. Actes del I Symposium Internacional de Lexicografía (Barcelona, 16-18 de maig de 2002)*, María Paz Battaner y Janet DeCesaris (eds), 591-604. Barcelona: Institut Universitari de Lingüística Aplicada – Universitat Pompeu Fabra.
- Montoro del Arco, Esteban T. 2005. "Hacia una sistematización de la variabilidad fraseológica". En *Estudios lingüísticos en recuerdo del profesor Juan Martínez Marín*, M^a Ángeles Pastor Milán (ed), 125-152. Granada: Universidad de Granada.
- Montoro del Arco, Esteban T. 2006. *Teoría fraseológica de las locuciones particulares. Las locuciones prepositivas, conjuntivas y marcadoras en español*. Frankfurt am Main: Peter Lang.
- Montoro del Arco, Esteban T. 2008. "El concepto de locución con casillas vacías". En *Colocaciones y fraseología en los diccionarios*, Carmen Mellado Blanco (ed), 131-146. Berlín/ Frankfurt am Main: Peter Lang.
- Moya Corral, Juan Antonio et al. 2007. *El español hablado en Granada. Corpus oral para su estudio sociolingüístico. I Nivel de estudios alto*. Granada: Universidad de Granada.
- Moya Corral, Juan Antonio et al. 2008. *El español hablado en Granada II. Corpus oral para su estudio sociolingüístico. Nivel de estudios medio*. Granada: Universidad de Granada.
- Moya Corral, Juan Antonio et al. 2009. *El español hablado en Granada III. Corpus oral para su estudio sociolingüístico. Nivel de estudios bajo*. Granada: Universidad de Granada.
- Real Academia Española. 2009. *Nueva gramática de la lengua española*. Madrid: Espasa.
- Ruiz Gurillo, Leonor. 1997. *Aspectos de fraseología teórica española*. Valencia: Universidad de Valencia, Departamento de Filología Española.
- Sosinski, Marcin. 2008. *Tratamiento de fraseología en los diccionarios bilingües polaco-español, español-polaco*. Granada: Universidad de Granada.
- Zuluaga Ospina, Alberto. 1980. *Introducción al estudio de las expresiones fijas*. Frankfurt: Peter Lang.