

El uso transitivo y ditransitivo de *dar* en un corpus escrito contrastivo

The Transitive and Ditransitive Use of *dar* in a Contrastive Written Corpus

Anna Sánchez Rufat
Universidad de Extremadura

Recibido: 24/07/2016
Aceptado: 17/10/2016

RESUMEN Los verbos frecuentes se consideran problemáticos para los aprendientes de lenguas extranjeras porque han desarrollado en cada lengua muchas especificidades. No obstante, escasean los trabajos empíricos que arrojen algo de luz sobre este aspecto de la interlengua del español. Este artículo analiza el uso del verbo *dar*, uno de los más representativos del grupo de verbos frecuentes del español, en la lengua escrita de un grupo de aprendientes anglófonos. En particular, se centra en las construcciones sintácticas más distintivas de este verbo, como son las transitivas y las ditransitivas. Esta investigación muestra la manera en la que un enfoque lingüístico basado en un corpus informatizado aplicado a los estudios de interlengua del español contribuye a revelar las semejanzas y diferencias entre el español de un grupo de hablantes no nativos y un grupo equivalente de hablantes nativos. El corpus utilizado para este análisis es el Corpus escrito del español como lengua segunda (Lozano, 2009). Las diferencias y similitudes entre el español de los nativos y el español de los no nativos permiten identificar las necesidades de estos referidas al verbo *dar* y determinar si estamos ante un verbo seguro (su uso no muestra desviaciones perceptibles) para el no nativo, o, por el contrario, constituye una fuente de problemas. Los resultados muestran que los aprendientes, incluso en un nivel avanzado, tienen dificultades con un verbo tan frecuente como *dar*. En las conclusiones, se subrayan algunas de las limitaciones de la investigación y se abren vías para futuros estudios que completarían el presente.

PALABRAS CLAVE: Análisis Contrastivo de la Interlengua; Investigación de Corpus de Aprendientes; interlengua del español; adquisición de ELE; verbo *dar*

ABSTRACT High frequency verbs tend to be problematic for foreign language learners, since they have developed many language-specific differences. However, there is a lack of empirical studies which shed light on this aspect of learners' Spanish interlanguage. This article explores Spanish learners' use of the verb *dar* (to give), one of the most representative verbs in the high frequency verbs group. In particular, it focuses on the most distinctive syntactic constructions of this verb, such as transitive and ditransitive usage in essay writing. This research shows how a corpus-linguistic approach to studies on learners' Spanish interlanguage reveals similarities and differences between the written Spanish of a group of non-native speakers and an equivalent group of native speakers. In turn, these similarities and differences allow us to identify learners' needs regarding their use of the verb *dar* in essay writing and to determine whether this is a safe verb (i.e. their use of it does not show noticeable deviations) for non-native speakers, or if it is a source of problems. The corpus used for this analysis is the Corpus escrito del español como lengua segunda (Lozano, 2009). The results show that learners, even at an advanced level, have difficulties with high frequency verbs such as *dar*. The conclusions highlight some of the limitations of this research and open up avenues for future research that would complete the present study.

KEY WORDS: Contrastive interlanguage analysis; Learner corpus research; Spanish interlanguage; Acquisition of ELE; the verb *dar*.

1. Los verbos frecuentes en el español como lengua extranjera. El caso del verbo *dar*

Todas las lenguas parecen tener un grupo de verbos que se utilizan continuamente en el discurso, por lo que ocupan los primeros puestos en las listas o diccionarios de frecuencias. El

Dirección de correspondencia:

Anna Sánchez Rufat, Universidad de Extremadura. Instituto de Español como Lengua Extranjera, Facultad de Filosofía y Letras. Avda. de la Universidad, s/n (Campus Universitario), 10003 Cáceres. E-mail: annasanchezrufat@hotmail.com. ORCID: 0000-0001-7125-1323.

verbo *dar* ocupa la décima posición entre los 1071 verbos registrados en *A Frequency Dictionary of Spanish*, de Davies (2006, p. 13), de 20 millones de palabras; y en De Miguel (2011, p. 141), en Lozano Zahonero (2010, pp. 88-89) y en la *Nueva Gramática de la Lengua Española. Manual* (2010, p. 670) es introducido como el verbo de apoyo más frecuente en español, el que tiene mayores posibilidades combinatorias. En esta misma línea, Herrero Ingelmo (2002, p. 6) presenta este verbo en el primer lugar del grupo de los verbos de apoyo de amplio espectro, es decir, aquellos que actualizan un mayor número de sustantivos predicativos; y en el corpus nativo utilizado para esta investigación, el Corpus Escrito del Español como Lengua Segunda o CEDEL2 (Lozano, 2009; Lozano y Mendikoetxea, 2013), el promedio de aparición del verbo *dar* es de una vez cada 546 palabras, lo cual supone una aparición por cada página de texto escrito.

El uso tan extendido de este verbo sugiere que los hablantes no nativos (HNN) deben adquirirlo. Parece evidente que el hecho de que un elemento sea frecuente en la lengua implica que el HNN está más expuesto a él y, por lo tanto, resulta más fácil de recordar (Nation, 2001; Nagini, 2008; Schmitt, 2010; Schmitt y Redwood, 2011). No obstante, las características atribuidas normalmente a los verbos frecuentes (Viberg, 1996) y nuestra experiencia en el aula nos inducen a pensar que los profesores y los alumnos los perciben como elementos difíciles de manejar. Esto ocurre por dos razones.

a) En primer lugar, la construcción verbo+nombre (V+N), que es –de todas las estructuras en las que puede aparecer el verbo *dar*– la más frecuente (presente en los usos ditransitivos y transitivos¹), y, aparentemente, la más sencilla por su estructura formal (compárese con las construcciones pronominales o preposicionales, entre otras), puede resultar compleja desde una perspectiva semántica. El contenido intensional de este verbo en esta construcción (V + N) es tan bajo que su capacidad extensional es máxima, esto es, puede coaparecer con una gran cantidad de sustantivos². En relación con ello, las restricciones que ejerce este verbo sobre sus argumentos son, muy a menudo, intralingüísticas, por lo que están basadas en la selección léxica del verbo en función de su estructura interna y de los rasgos subléxicos contenidos en las palabras con las que se combina y con los que concuerda; y estos rasgos fácilmente difieren de una lengua a otra. Por ello, aunque en español tenemos *dar un libro a alguien* y en inglés *to give a book to somebody* (la restricción extralingüística del verbo –basada en nociones físicas– es la misma³), y aunque también tenemos *dar cariño a alguien* y *to give love to somebody* (ejemplos de restricción intralingüística compartida), la combinatoria de ambos verbos (*dar* y *to give*) no es, en absoluto, equivalente en muchos casos.

b) En segundo lugar, y en relación con lo anterior, el verbo *dar* y el verbo *to give* difieren de manera importante en las construcciones con verbo de apoyo (CVA), como en *dar un salto* –cuya característica semántica distintiva es el significado abstracto del verbo, próximo a “hacer” (Bosque, 2001, p. 23) o “ejecutar”–, por lo que estas estructuras pueden ser un foco de desviaciones para los HNN. Trabajos como el de Wanner *et al.* (2013) refuerzan esta hipótesis en tanto en cuanto sus conclusiones sobre las colocaciones muestran que los HNN cometen más errores en las CVA, que usan relativamente a menudo, que en otras combinaciones V+N.

Las CVA son opacas –comparadas con las construcciones en las que *dar* actualiza su significado pleno (CVP), próximo a “proporcionar”, como en *dar cariño*–, ya que el vínculo entre el significado figurado del verbo de apoyo –“hacer”– y el significado básico o físico a partir del cual se proyecta –“entregar una entidad material”, como en *dar un lápiz*– es difícil de rastrear en ellas. Así, los sentidos figurados se proyectan a partir del significado básico, pero este no está tan visible en *dar un salto* o un *paseo* (CVA) como en *dar cariño* o una *opinión* (CVP).

¹ Ambas estructuras (transitivas y ditransitivas) comparten la transitividad; lo único que las diferencia es la presencia explícita del complemento indirecto en la ditransitivas.

² Para una reflexión sobre la naturaleza léxico-semántica y la proyección sintáctica del verbo *dar*, véase Sánchez Rufat (2015a).

³ Para una explicación más detallada de los conceptos *restricción intralingüística* y *extralingüística*, véase Sánchez Rufat (2010 y 2011).

Consideramos que tras el significado básico y los significados figurados (ya sea el que actualiza el verbo en las CVA o en las CVP) se oculta un mismo significado *sistémico* (al que se refería Rhul [1989] con el término *inherente*) a partir del cual todos los usos del verbo se proyectan (básico y figurados), que en el caso de *dar* sería: (A) causa que B pase de no existir a existir (en C, de manera que los seres C disfruten o padezcan B); pudiendo ser B *un lápiz, un salto o una opinión*. Así pues, todos los significados del verbo están contemplados en la definición sistémica; no obstante, el significado que se percibe en las CVA es fundamentalmente el aspectual, dado que la gramática en estas construcciones prescinde de la idiosincrasia léxica del verbo –que en el caso de *dar* es la transferencia– y selecciona fundamentalmente los rasgos aspectuales del significado del verbo –próximo a “hacer, ejecutar o producir”– (Bosque, 2001). Esto provoca ese distanciamiento entre el significado figurado del verbo de apoyo y el significado básico, físico, a partir del cual se proyecta; lo que revierte en la dificultad a la hora de descodificar y, principalmente, codificar estos usos lingüísticos. Asimismo, el hecho de que dicho significado aspectual sea compartido por otros verbos de apoyo (como *hacer*) contribuye, de nuevo, a esa percepción de arbitrariedad y a que no haya necesariamente una coincidencia en las gramáticas de las diferentes lenguas a la hora de seleccionar un verbo de apoyo para un sustantivo eventivo, como son *salto o paseo*⁴.

Todo esto hace que posiblemente la L1 sea una fuente de interferencias en estas CVA –mientras que en español *damos un paseo o un paso*, en inglés se *toman (take a walk y take a step)*– o que el verbo de apoyo sea elegido al azar de entre varios verbos frecuentes (Chi Man-lai *et al.*, 1994), dado que, como se ha señalado, al codificar un mensaje, el HNN no interpreta con la misma facilidad la extensión natural del significado básico al significado figurado en las CVP (*dar cariño o una opinión*) que en las CVA (*dar un salto o paseo*).

En suma, pese a las semejanzas semánticas entre las lenguas, los verbos frecuentes han desarrollado en cada lengua muchas especificidades (Altenberg y Granger, 2001, p. 174). Ténganse en cuenta, por tanto, estas diferencias interlingüísticas y complejidades intralingüísticas –semánticas– que, a nuestro juicio, pueden dificultar la adquisición del verbo *dar*. Las hipótesis formuladas en torno a la interfaz léxico-sintaxis acerca de los posibles focos de desviaciones en el uso de este verbo tan extendido constituyen un certero punto de partida para el análisis que aquí se propone, motivado no solo por nuestra experiencia en el aula y la escasez de trabajos que arrojen algo de luz sobre este aspecto de la interlengua del español (para un ejemplo de los pocos estudios empíricos dedicado a verbos frecuentes del español, véase Molina Alfaro, 2013), sino también por las discrepancias observadas en la bibliografía revisada en torno al uso de los verbos frecuentes en inglés por parte de grupos de aprendientes de diferentes L1. Hay estudios que se refieren a la sobreutilización de estos verbos por parte de los aprendientes (Hasselgren, 1994; Granger, 1996a; Ringbom, 1998; Källvist, 1999; Guo, 2006), mientras que otros hablan de infrautilización, como Altenberg y Granger (2001) o Sinclair (1991, p. 79): “many learners avoid the common verbs as much as possible, and especially where they make up idiomatic phrases. Instead of using them, they rely on larger, rarer, and clumsier words which make their language sound stilted and awkward”.

Ahora bien, para identificar qué aspectos de la interlengua del español, en lo referido a los usos transitivos y ditransitivos de *dar*, son producidos en consonancia con los usos nativos, cuáles se distancian un tanto y cuáles son los que se desvían claramente del español nativo, no nos basta con conocer las semejanzas y diferencias en cuanto a las frecuencias de uso, sino que también se requiere un análisis detallado de la actuación, que va a permitir, además, determinar la adecuación de cada uso de acuerdo con el concepto de norma prefijado. Y a este respecto tampoco parece haber acuerdo entre los investigadores de la interlengua del inglés. Mientras que algunos establecen que estamos ante verbos “seguros” para el HNN –como señala Hasselgren (1994, p. 250) en relación con su propuesta *teddy bear principle*, según la cual “core words, learnt early, widely usable, and above all safe (because they do not show up as errors) are hugely overused, even among

⁴ En Sánchez Rufat (2015a) se explica detenidamente la relación entre el sentido básico o físico del verbo y los sentidos figurados, así como la relación existente entre ambos tipos de significado y el significado sistémico, a partir del cual los anteriores se proyectan.

learners sufficiently advanced to have been weaned off them”—, otros consideran que son una fuente de errores (Chi *et al.*, 1994; Howarth, 1996; Källvist, 1999; Altenberg y Granger, 2001; Gui y Yan, 2001; Guo, 2006). Esta falta de acuerdo entre los investigadores de la interlengua del inglés acerca de si estos verbos se sobreutilizan o infrautilizan, o si son seguros o una fuente de errores, unido, a su vez, a la escasez de trabajos que se centran en este aspecto de la interlengua del español, refuerzan, por lo tanto, la necesidad del presente análisis.

Teniendo en cuenta estas consideraciones, esta investigación arranca de las siguientes preguntas, cuyas respuestas permitirán confirmar o descartar las hipótesis anteriores:

1. ¿Cuáles son las semejanzas y diferencias más destacadas entre el español de los HNN y el español de los hablantes nativos (HN) con respecto al uso del verbo *dar* en las construcciones transitivas y ditransitivas?

1.1. ¿Tienden los HNN a utilizar este lema más o menos que los HN?

1.2. ¿Cuál es la relación del índice de frecuencias en cada uno de los dos tipos de construcción sintáctica?

1.3. ¿Con qué palabras o tipos de palabras combinan los HN y los HNN el verbo *dar*?

2. ¿Es un verbo *seguro* (su uso no muestra errores perceptibles) para el HNN o, por el contrario, es una fuente de problemas fruto de las diferencias interlingüísticas y complejidades intralingüísticas?

3. ¿Qué papel desempeña la lengua materna (L1) en el uso de este verbo? y ¿qué presencia tiene el factor intralingüístico en la adquisición de este verbo?

Para contestar a estas preguntas, comparamos datos de lengua de aprendientes y datos de HN por medio de un corpus informatizado que agiliza la fase inicial del análisis.

2. Datos y metodología

Esta investigación pertenece al campo de estudio de la Adquisición de Segundas Lenguas (ASL); en concreto, forma parte de la rama de estudios de la lengua del aprendiente. Se aplica para el análisis un enfoque basado en corpus a través de la metodología específica del Análisis Contrastivo de la Interlengua o ACI (Granger, 1996b y 2015), por medio de la cual se estudia el español escrito de un grupo concreto de aprendientes de nivel avanzado de L1 inglés –el español de los HNN del CEDEL2– y se compara con el español nativo de un corpus de control, también integrado en el corpus CEDEL2; esta comparación permite descubrir las semejanzas y las diferencias entre los usos nativos y los no nativos.

El CEDEL2 nace en 2006 en el grupo de investigación WOSLAC (Word Order in Second Language Acquisition Corpora), de la Universidad Autónoma de Madrid y la Universidad de Granada, y constituye una fuente de datos reales y naturales de calidad metodológica al haber sido creado de acuerdo con los diez principios del diseño de corpus propuesto por Sinclair (2005). Es un corpus escrito informatizado no lematizado de HNN de L1 inglés-L2 español que contiene muestras de lengua de todos los niveles de dominio (inicial, intermedio y avanzado) y que se encuentra todavía en la fase de recopilación de datos: se pretende llegar al millón de palabras (250000 de cada nivel); en este estudio se han utilizado 198000 palabras nativas y 172000 no nativas de nivel avanzado, que eran las muestras disponibles en el momento en que realizamos este análisis. De acuerdo con los especialistas en este campo, estas cantidades son consideradas suficientes para aportar un panorama fiable de cualquier uso lingüístico (Lozano y Mendikoetxea, 2013, p. 16). Los HNN son clasificados de acuerdo con el test de diagnóstico de la Universidad de Wisconsin (1998), que es independiente y estandarizado, y se puede realizar en línea; y, al mismo tiempo, el CEDEL2 dispone de un corpus nativo diseñado de igual manera con el objetivo de que sirva para establecer comparaciones entre la lengua de los HNN y la de los HN. Los textos corresponden a HNN que han adquirido la LE en un contexto de aprendizaje formal, para cuya producción no ha habido ninguna restricción en el contenido lingüístico ni temático; por eso, el corpus se considera un corpus representativo del estado de lengua elegido para este análisis: los HNN participaron

voluntariamente y pudieron elegir libremente un tema de redacción de entre 12 propuestos, unos descriptivos y otros más argumentativos –por lo que las muestras de lengua son variadas y no reflejan las restricciones léxicas de un solo tipo de texto (Guo [2006: 55] se refiere a que “different genres or text types will involve a different lexis”; véase Lozano y Mendikoetxea [2013: 9] para algunos clichés léxicos propios del ensayo argumentativo)–. Asimismo, dado el carácter neutro de los tipos de texto utilizados (redacciones de clase), es fácil que los problemas encontrados se reproduzcan en otros tipos de texto, tal y como argumenta Nesselhauf (2005, p. 46) a propósito de su análisis de colocaciones basado en corpus: “as the text type investigated is highly neutral, i.e. not specialised and on a medium level of formality, the difficulties identified are likely to occur in most other text types as well”.

El CEDEL2 contiene información precisa y detallada de cada estudiante y de las condiciones en las que es elaborada la redacción, lo que es esencial en una investigación sobre la lengua del aprendiente, pues permite filtrar los datos y descartar los textos que no cumplen los criterios del corpus, así como indagar sobre las posibles causas de los resultados “extraños” o “llamativos”.

Dado que no debemos asumir que el español de los HN puede erigirse en la única norma para los HNN (presenta problemas y ciertas marcas de distinto tipo), es necesaria la aplicación de una norma reguladora que precise el grado de adecuación de los usos no nativos. Para un corpus como el CEDEL2, en el que tanto los HN como los HNN manejan referentes diatópicos distintos, la unión de los estándares peninsular y americano en un estándar panhispánico constituye un referente lingüístico operativo para categorizar los datos de la interlengua (abordamos la cuestión de la norma en el análisis de la interlengua en Sánchez Rufat y Jiménez Calderón, 2013). En los casos en los que se alberguen dudas sobre la estandarización de algún elemento y se sienta la necesidad de contrastar las intuiciones, se recurre a herramientas como un corpus más amplio (CREA), diccionarios apropiados (*Redes, Práctico, Diccionario panhispánico de dudas*, etcétera), la *Nueva gramática de la lengua española* o encuestas a informantes nativos de diferentes variedades diatópicas, tanto peninsulares como americanas (dos procedentes de Cádiz y Salamanca, y otros dos procedentes de Argentina y México), los cuales han pasado toda su vida en un país de habla hispana y cuyo perfil coincide con el de los autores del corpus: son universitarios o graduados. Debido a que el número de informantes es bajo, los juicios de aceptabilidad han de considerarse como una aproximación, juzgada como suficiente para los objetivos marcados en este estudio. No obstante, hay trabajos que muestran que la correlación entre los datos del corpus y los juicios de los nativos en coapariciones léxicas es buena (Lapata *et al.*, 1999; Shei, 1999; Hoffmann y Lehmann, 2000). De hecho, antes de enviar los datos a los cuatro informantes nativos, se ha realizado una prueba piloto con varias combinaciones que demuestra acuerdo entre los nativos en aquellas combinaciones inexistentes, y por tanto, inaceptables, y en aquellas frecuentes en el mundo panhispánico, y por tanto, aceptables-naturales.

El *software* utilizado para la recuperación de los usos de *dar* es Antconc3.2.3m y el método para analizar los datos obtenidos a partir de las concordancias es cuantitativo y cualitativo. A este respecto, Nesselhauf (2004, p. 136) señala lo siguiente:

Many studies are exclusively or primarily quantitative...While such studies can be interesting starting points for further quantitative analyses, they do not usually in themselves contribute much to language learner analysis, let alone to language teaching. If progress is to be made, it is imperative that this current stage is left behind and that more qualitative analyses are carried out.

Por tanto, las reflexiones teóricas expuestas en el análisis surgen de los resultados estadísticos obtenidos de la investigación cuantitativa. Solo así es posible aprovechar las ventajas de la tecnología informática actual y la práctica y las teorías de la Lingüística de Corpus y de otras áreas de la ASL, como la Psicolingüística y la enseñanza de L2.

3. Frecuencia total de *dar*

El primer paso del análisis consiste en averiguar si este grupo de HNN de ELE utiliza el verbo *dar* con la misma frecuencia que los HN. Ninguna estructura en la que aparece este verbo ha sido descartada (transitiva, ditransitiva, intransitivizada, impersonal, pronominal, pasiva, construcción de participio absoluto, construcción preposicional, locución y esquema fraseológico semiproductivo). Para determinar la frecuencia de uso del lema en ambos corpus, recurrimos, en primer lugar, al programa Antconc (versión 3.2.3.m) para obtener todas las ocurrencias a través de las concordancias. Como ya se ha señalado, el CEDEL2 no es un corpus lematizado en el momento en que realizamos esta investigación, por lo que para poder localizar en él todas las formas del paradigma verbal se han introducido en la búsqueda todas las posibilidades de flexión verbal. Al mismo tiempo, se ha atendido a los posibles errores ortográficos en los que hayan podido incurrir los HNN al escribir el verbo, para introducirlos también en la búsqueda. No obstante, no es posible descartar que algunas formas ortográficas erróneas de *dar* no hayan sido capturadas. A continuación, se han excluido de las concordancias obtenidas todos los datos irrelevantes para este estudio (como los nombres propios, el sustantivo *dado* o la conjunción *dado que*). En números absolutos, los resultados obtenidos son 360 usos del lema en los HN y 218 en los HNN. Tras obtener estos datos, contamos las frecuencias normalizadas (ocurrencias por 100000 palabras) del lema *dar* en los dos corpus⁵. Los resultados se recogen en la tabla 1:

Tabla 1
Frecuencia de *dar* en HN y HNN

| Lema | HNN | HN |
|------------|---------|-------|
| <i>Dar</i> | 127, 6* | 183,1 |

Por cada 100000 palabras, los HN emplean *dar* en 183,1 ocasiones, frente a las 127,1 de los HNN. Así, por cada ocurrencia de *dar* en el corpus no nativo, el lema ocurre 1,4 veces en el nativo. La tabla muestra, por tanto, una diferencia estadísticamente significativa ($p < 0,05$), marcada en las tablas por medio de un asterisco tras la cifra correspondiente, entre la frecuencia de uso de *dar* de los HNN y los HN; los HNN usan el lema *dar* menos que los HN.

A continuación, profundizamos en los usos transitivos y ditransitivos de *dar*, considerando que este verbo expresa una amplia variedad de significados y se combina con una gran cantidad de sustantivos⁶.

4. Frecuencia de *dar* en las construcciones transitivas y ditransitivas

La tabla 2 recoge los dos tipos de estructuras analizadas e ilustra cada una de ellas con un ejemplo del corpus.

Tabla 2
Tipos de construcciones sintácticas en las que participa *dar*

| Estructuras | Ejemplos |
|-----------------|-------------------------------------|
| 1. Transitiva | Algunas pistas dan mucho miedo |
| 2. Ditransitiva | Los fuegos me dan un poco de reparo |

⁵ A partir de la suma de las apariciones del lema *dar* en el corpus (frecuencia absoluta) se calcula su *porcentaje* de uso. Dado que los dos corpus no son exactamente iguales en tamaño –el corpus nativo cuenta con 196623 palabras y el no nativo con 170849 palabras, por lo que el primero excede en unas 25000 palabras al segundo–, necesitan normalizarse. En esta ocasión el porcentaje (%) no transmite la frecuencia de uso de la palabra de manera significativa (0,00183 %), por lo que se calcula una *frecuencia normalizada* a partir de una base de normalización comparable con la medida del corpus, como es 100000.

⁶ Para un análisis exhaustivo de todas las construcciones sintácticas proyectadas por el verbo *dar* en el corpus CEDEL2, véase Sánchez Rufat (2015b). En el presente trabajo nos ceñimos únicamente a los usos transitivos y ditransitivos del verbo.

En la tabla 3 se recogen las frecuencias de *dar* por cada construcción sintáctica. RF representa el recuento de frecuencia o frecuencia absoluta y FN representa la frecuencia normalizada. El porcentaje (%) se refiere a la frecuencia absoluta dividida por la frecuencia total de *dar* en el corpus correspondiente.

Tabla 3

Frecuencia de dar en las diferentes estructuras

| Estructura | NoNativos | | | Nativos | | |
|---|-----------|-------|-------------|---------|------|-------------|
| | RF | FN | % | RF | FN | % |
| Transitiva | 39 | 22,9 | 17,8 | 62 | 31,5 | 17,2 |
| Intransitiva | 5 | 2,9 | 2,3 | 14 | 7,1 | 3,9 |
| Ditransitiva* | 106 | 62 | 48,4 | 101 | 55,9 | 30,6 |
| Impersonal | 12 | 7 | 5,9 | 30 | 15,3 | 8,3 |
| Pasiva | 3 | 1,7 | 1,4 | 8 | 4 | 2,2 |
| C. Absoluta | 3 | 1,7 | 1,4 | 5 | 2,5 | 1,4 |
| Pronominal | 1 | 0,6 | 0,5 | 12 | 6,1 | 3,3 |
| C. preposicional | 2 | 1,2 | 0,9 | 12 | 6,1 | 3,3 |
| Construcción con <i>como</i> | 0 | 0 | 0 | 3 | 1,5 | 0,8 |
| Adjetivo | 0 | 0 | 0 | 3 | 1,5 | 0,8 |
| Locución | 45 | 26,3 | 20,5 | 93 | 47,4 | 25,8 |
| Esquema fraseológico semiproductivo | 2 | 1,2 | 0,9 | 8 | 4 | 2,2 |
| Total | 218 | 127,5 | 100 | 360 | 183 | 100 |

Los resultados, por un lado, ponen de manifiesto que el *ranking* de usos de las dos estructuras es el mismo en los dos corpus: la estructura ditransitiva es la más utilizada, mientras que la transitiva ocupa la tercera posición. A simple vista, parece que el análisis de frecuencias indica que el uso de *dar* en los dos corpus es similar, pues ambas construcciones se producen con una frecuencia alta (superan el 5 % de aparición fijado⁷). La alta frecuencia de uso de estas estructuras en el corpus no nativo no resulta sorprendente, dado que en el input al que están expuestos los HNN estas construcciones son altamente recurrentes y, al mismo tiempo, estas estructuras, desde una perspectiva formal, no son complejas y sintácticamente son equivalentes en la L1; además, caracterizan el léxico restringido de la oralidad que es propio de los HNN en tareas escritas de tema libre (Cobb, 2003). No obstante, si en lugar de centrarnos en el *ranking* profundizamos en los índices de frecuencias por medio de la aplicación del test de significatividad Chi-cuadrado, se observan importantes diferencias entre un grupo y otro.

En primer lugar, se observa que los HNN pueden al mismo tiempo utilizar significativamente menos un verbo frecuente y utilizar significativamente más ciertas estructuras, como se desprende de los datos de las ditransitivas, que son empleadas por los HNN un 14 % más. Estas estructuras son frecuentes en ambos corpus, pero hay una diferencia en la frecuencia que es estadísticamente significativa ($p < 0,05$). En cambio, las estructuras transitivas, también frecuentes, se usan en ambos grupos por igual, en un 17 % de los casos. Como se describe en un apartado posterior (5.2), la sobreutilización de las ditransitivas radica, en parte, en el hecho de que los HNN sobreutilizan las combinaciones de *dar* con las clases de sustantivos que denotan “sensaciones físicas, psíquicas o anímicas” (*ataque de nervios, alegría, felicidad, miedo*). En estas estructuras siempre existe un experimentador de dichas emociones, que es el complemento indirecto (CI); este no está

⁷ Un porcentaje del 5 % o más se considerará frecuencia alta, y uno menor que el 5% se considerará frecuencia baja. Aunque este límite es arbitrario, está basado en nuestra experiencia y coincide con el marcado por Guo (2006, p. 182), quien trabaja con un corpus de tamaño muy similar al nuestro.

presente en las estructuras transitivas, lo cual explica que el sobreuso no las afecte. Asimismo, el hecho de que las ditransitivas se sobreutilicen puede sugerir que se desconocen otros recursos expresivos más adecuados en ciertos contextos, como sucede en aquellos casos en los que se debería o podría emplear un solo verbo en lugar del verbo *dar* + sustantivo deverbal. Así, en usos como estos⁸ puede esconderse otra causa de la sobreutilización:

[1] Me da gusto saber que puedo ayudarles a mejorar sus vida. En lugar de *Me gusta saber*.

[2] el gobierno dice que [el cannabis] es "una droga" que actúa como una entrada a las drogas mas duras. (...) Es una mentira del gobierno para dar miedo a los jóvenes. En lugar de *asustar*.

[3] [en Año Nuevo] Emery yo nos dimos besos a media noche porque es la tradición. En lugar de *nos besamos*.

5. Palabras y tipos de palabras con los que se combina *dar*

Ya hemos señalado que el verbo *dar* ocupa una de las primeras posiciones en la relación de los verbos más frecuentes del español por la asiduidad con la que se utiliza la construcción V + N, que se corresponde con las estructuras en las que el verbo actualiza los sentidos de “proporcionar” y “hacer, ejecutar o producir”. En ellas, el nombre es su argumento interno y tema; el tema es el complemento directo (CD) en las estructuras transitivas y ditransitivas.

Para poder codificar correctamente estas estructuras, no basta con saber que el verbo exige un CD, sino que también se ha de conocer el contenido semántico que desempeña la función semántica tema. Dado que este verbo es el que tiene mayores posibilidades combinatorias (Lozano Zahonero, 2011, p. 89), el nombre que lo acompaña puede pertenecer a muchas clases léxicas. De este modo, los predicados seleccionan sus argumentos restringiendo el conjunto de piezas léxicas que pueden denotar de acuerdo con sus rasgos semánticos, por lo que las unidades seleccionadas no son piezas aisladas, sino nociones muy diversas que pueden ser denotadas por grupos de sustantivos que son seleccionados por el verbo. Así, *dar un paseo* no constituye una combinación en la que la relación entre los elementos es biunívoca, sino que es un caso de selección léxica por medio de la cual el verbo selecciona los sustantivos que denotan “desplazamiento, más frecuentemente si es breve o repentino y se interpreta figuradamente”, como *vuelta, salto, paso, vuelco, rodeo, giro*, etcétera (en el diccionario combinatorio *Redes* se localizan las clases léxicas seleccionadas por este verbo).

5.1. El uso de *dar* como verbo de transferencia

Los HN emplean *dar* con el sentido de “proporcionar” con 20 clases léxicas, y los HNN, con 19. La mayoría de las clases son usadas por ambos grupos con una frecuencia semejante (diferencia menor del 5 %), pero hay dos que utilizan más los HN y otras dos que utilizan más los HNN. En esta tabla presentamos las clases léxicas que se usan con una frecuencia diferente en un 5 % o más. Entre corchetes se localizan los sustantivos utilizados en el corpus. El porcentaje entre paréntesis se refiere a la diferencia que se da entre los porcentajes de ambos grupos, y el asterisco representa la significatividad estadística ($p < 0,05$):

⁸Todas las muestras de lengua incluidas a partir de ahora proceden de los datos del CEDEL2.

Tabla 4

Clases léxicas seleccionadas por dar cuando actualiza el sentido de “proporcionar” (en las CVP) con diferencias significativas entre el uso nativo y no nativo

| + Frecuencia HN | + Frecuencia HNN |
|--|---|
| Asistencia, sustento o protección [<i>apoyo, consejo, beca, protección, habitación, atención médica, abrigo, amparo</i>] (12,5 %)* | Ocasión, oportunidad o situación favorable [<i>oportunidad(es), opción, ventaja</i>] (19,4 %)* |
| Valores y cualidades de las personas o cosas [<i>atractivo, educación, color, realismo, autenticidad, prestigio</i>] (7 %)* | Entidades materiales [<i>billete, tenedores, obras, mate, bebidas, comida, peluche, regalos, medicinas, bocadillos</i>] (5 %) |

Esta tabla ilustra la razón de una pequeña parte de esa diferencia que existe en la frecuencia de uso de las estructuras V + N entre ambos grupos, que, como ya se ha señalado, es de un 14 % (los HNN las usan más). Llama la atención la alta frecuencia con la que los HNN utilizan la combinación *dar una oportunidad* frente al bajo porcentaje de uso de los HN (19 frente a los 3 usos nativos). Dada la baja frecuencia de este uso en el corpus nativo, no podemos estar seguros acerca de la significatividad de su alta presencia en el corpus no nativo. Cuando las concordancias se examinan, la mayoría de ellas son aceptables. Así, parece que esta coaparición disfruta de una popularidad y homogeneidad entre este grupo de HNN.

Para concluir, el uso productivo del verbo *dar* en lo que se refiere a los sustantivos con los que se combina en la construcción V + N cuando el verbo actualiza el significado de “proporcionar” es prácticamente igual de amplio: el 55,2 % en los HN y el 58,1 % en los HNN de las combinaciones de *dar* son diferentes en cada corpus; el porcentaje restante representa usos repetidos de una misma combinación. Pese a este porcentaje similar, los sustantivos con los que cada grupo combina *dar* no coinciden en la mayoría de casos (86 %), y, al mismo tiempo, se dan diferencias significativas en el porcentaje de uso de las cuatro clases léxicas señaladas o en el de algunas combinaciones léxicas –como en el sobreuso de *dar la oportunidad*–.

5.2. El uso de *dar* como verbo de creación

Cuando el verbo *dar* se utiliza como verbo de apoyo, esto es, con el sentido de “hacer o producir”, hemos distinguido dos usos del verbo muy próximos – basándonos en la organización combinatoria de *dar* en *Redes* (Bosque, 2005, pp. 639-641)–: cuando se usa como *producir* en el sentido de “ejecutar” y cuando se utiliza como *producir* en el sentido de “suscitar(se) en alguien”. Así, el primer uso se corresponde con el de *Le dieron una patada*, y el segundo con el de *Le dieron miedo*.

Con el sentido de “ejecutar”, *dar* se combina con sustantivos que designan “movimientos resultantes de algún contacto” (*abrazo, empujón, toque*), sustantivos que denotan “comienzo y final”, sustantivos que denotan “desenlace y nociones relativas al fin de las acciones o los procesos” (*respuesta, solución, cumplimiento*), sustantivos que denotan “enfrentamiento” (*batalla, pelea*); los que denotan “molestia” (*lata, tostón, tabarra*); los que se interpretan como “unidades informativas y, por extensión, los que expresan eventos que se ofrecen a la contemplación pública” (*clase, lección, concierto, rueda de prensa*); los sustantivos que designan “formas de trato social y otros que se refieren a los sacramentos y a diversos signos que forman parte de ceremonias o actos religiosos” (*buenos días, bienvenida, pésame*); los sustantivos que denotan “desplazamiento breve” (*paseo, salto, paso*); y, por último, los que designan “sonidos” (*grito, silbido*).

Con el sentido de “suscitar(se) en alguien”, *dar* se combina con los sustantivos *sensación* e *impresión*, así como con varios otros que designan “alteraciones y reacciones físicas adversas, repentinas” (*calambre, mareo, retortijón*); sustantivos que denotan “deseo o necesidad” (*gana, hambre, envidia*); los que designan “sensaciones y emociones percibidas como placenteras” (*alegría, gusto, felicidad*); los que denotan “miedo, zozobra y estados anímicos similares” (*miedo, susto, terror, apuro*); y

los que designan “otros estados de aflicción relacionados con la infelicidad, el enojo o el rechazo” (*asco, tristeza, pena*).

Partiendo de esta información, los HN usan *dar* en las CVA con el sentido de “ejecutar” tres veces más que con el sentido de “suscitar(se)”, lo cual supone el 76 % frente al 24 %. En cambio, los HNN usan los dos sentidos del verbo prácticamente por igual, un 50,8 % de las ocasiones con el de “ejecutar” y un 49,2 % con el de “suscitar(se)”. Esta diferencia entre el uso nativo y no nativo es estadísticamente significativa ($p < 0,05$). Si profundizamos en el uso que cada grupo hace de las clases léxicas, observamos que la mayoría de las clases léxicas más utilizadas por ambos grupos son empleadas con una frecuencia semejante (diferencia menor del 5 %); pero de nuevo hay dos que utilizan más los HN y otras dos que utilizan más los HNN. En esta tabla presentamos las clases léxicas que se usan con una frecuencia diferente a un 5 %. Entre corchetes se localizan los sustantivos utilizados en el corpus. El porcentaje entre paréntesis se refiere a la diferencia que se da entre los porcentajes de ambos grupos. La diferencia es únicamente significativa desde una perspectiva estadística ($p < 0,05$) en la sobreutilización de los HNN:

Tabla 5

Clases léxicas con las que se combina *dar* cuando actualiza el sentido de “producir(se)” (CVA) con diferencias significativas entre el uso nativo y no nativo

| + Frecuencia HN | + Frecuencia HNN |
|--|--|
| Desplazamiento, frecuentemente breve y repentino [<i>vuelta, paso, paseo, salto, giro, voltereta</i>] (7,8 %) | Miedo, inquietud, zozobra y estados anímicos similares [<i>vergüenza, miedo, sorpresa</i>] (12,7 %)* |
| Movimientos resultantes de algún contacto [<i>vuelta de tuerca, patada, golpe, paliza, trompaço, beso, empujón, caricia</i>] (6,3 %) | Sensaciones y emociones placenteros [<i>gusto, placer, alegría, risa, satisfacción</i>] (10,1 %)* |

La mencionada sobreutilización del sentido de “suscitar(se)” (49,2 % de los usos no nativos de *dar* en las CVA frente al 24 % de los usos nativos) y la consiguiente infrautilización del de “ejecutar” pueden explicarse si se tiene en cuenta que los HNN, partiendo de géneros textuales equivalentes, muestran más implicación personal que los HN (Cobb, 2003); así, esta sobreutilización de *dar* con el sentido de “suscitar(se)” se traduce en frecuentes referencias al estado mental, físico y anímico del autor de la redacción, como se refleja en la tabla 5. De este modo, esta tabla y la tabla 4 muestran las causas de gran parte de esa diferencia del 14 % que existe en la frecuencia de uso de las estructuras V + N en ambos grupos. Vemos, pues, que gran parte de esa sobreutilización de la construcción V + N se debe a la tendencia de los HNN de implicarse personalmente en la narración, de referirse a su estado anímico.

Por otro lado, llama la atención que, precisamente en las dos clases léxicas más frecuentes entre los HNN –la de los “desplazamientos” y la de los sustantivos que denotan “miedo o zozobra”–, la variedad de sustantivos utilizados es menor que en otras clases léxicas menos frecuentes. Los HNN usan menos sustantivos en estas clases pero los repiten más; así, mientras que los HN usan *miedo, reparo* o *pánico*, los HNN, dentro de esta gama significativa, solo usan *miedo*; y mientras los HN *dan paseos, vueltas, giros, pasos* y *saltos*, usados en la mayoría de las ocasiones en sentido figurado (*dar un paso en la vida, un giro a la historia, vueltas a la cabeza, dar el gran salto, etcétera*), los HNN, en cambio, solo *dan vueltas* y *paseos* en sentido literal.

Cabe relacionar el menor empleo que hacen los HNN de esta última clase léxica (la de los “desplazamientos”) –pese a ser precisamente la más empleada en el corpus, tanto no nativo como nativo– con el hecho de que lo que convierte esta combinatoria en la más frecuente del corpus nativo sea el uso de los sustantivos correspondientes en sentido figurado; este conocimiento es muy específico, más resbaladizo que el de los sentidos literales, y es posible que la mayoría de los HNN no sea muy consciente del empleo figurado de sustantivos como *paso, salto, giro* o *vuelta*; y, por otro

lado, el inglés selecciona para esta clase léxica el verbo *to make*, principalmente, lo que también puede contribuir a la infrautilización de *dar* en este contexto.

Por lo tanto, parece que, si profundizamos dentro de una clase léxica, tras las aparentes semejanzas de uso se perfilan ciertas diferencias, fundamentadas en la presencia frente a la ausencia de sustantivos cuasisinónimos (*reparo, miedo, pánico*) y en la ausencia frente a la presencia de usos figurados.

Por otro lado, pese a las diferencias significativas en el porcentaje de uso de algunas clases semánticas, el 47,9 % de las combinaciones de *dar* en el caso de los HN y el 48,4 % en el de los HNN son diferentes; el porcentaje restante se refiere a usos repetidos de una misma combinación, por lo que el uso productivo de *dar* en lo que respecta a los sustantivos con los que se combina el verbo en la construcción V + N cuando el verbo actualiza el significado de “producir(se)” es prácticamente igual de amplio, si bien los sustantivos con los que cada grupo combina *dar* no coincide en la mayoría de los casos (77 %). Resulta llamativo que, de los diferentes sustantivos utilizados por cada grupo, el 87 % de los que se localizan en el corpus nativo pertenecen a las 5000 palabras más frecuentes del español, según la información de *A Frequency Dictionary of Spanish* (Davies 2006), mientras que el grupo de HNN utiliza hasta un 94,1 % de palabras frecuentes. Estos datos sugieren que, en línea con el estudio de Cobb (2003, p. 403), los aprendientes, cuando han de realizar una tarea escrita sobre un tema para un lector anónimo, utilizan el vocabulario más frecuente, más propio de la oralidad.

6. Aceptabilidad e inaceptabilidad de los datos

La frecuencia en el corpus de HNN solo revela cuántas veces ocurre un elemento determinado en el corpus. Aunque puede indicar cierto grado de adquisición en la producción del español por los HNN, no refleja necesariamente dominio, pues la lengua del aprendiente está cargada de rasgos problemáticos y expresiones artificiales. Por lo tanto, tras la frecuencia se ocultan los errores.

A continuación, presentamos la frecuencia de los errores y de los aciertos, en general, y ofrecemos el porcentaje de error y de acierto en las dos estructuras analizadas para este trabajo. Se señala el índice de frecuencias al respecto de la naturaleza –léxica o computacional– del error en el verbo. Estos datos de frecuencia están respaldados por un análisis detallado de las concordancias.

6.1. Frecuencia de los datos aceptable-naturales, aceptable-restringidos⁹ y de los inaceptables en el uso general del verbo *dar*

El análisis revela que, pese a que los índices de frecuencia generales no son drásticamente diferentes, tras ellos a menudo se esconden errores. No podemos decir que el verbo *dar* sea un verbo seguro para este grupo de HNN, dado que se comenten errores en 1/3 de los usos; de estos, más de la mitad se produce por un desconocimiento del verbo (56,3 %), frente a los errores originados en alguno de los argumentos del verbo. En el siguiente gráfico se representan estas frecuencias:

⁹ Los usos clasificados como naturales no difieren en nada del estándar, esto es, no presentan ninguna desviación ortográfica, semántica, gramatical, combinatoria, diatópica, diastrática ni diafásica; en cambio, los usos restringidos son muestras de usos correctos desde una perspectiva formal, pero considerados no preferentes por los informantes nativos a los que se ha recurrido en este trabajo. El término *inaceptabilidad* y otros términos relacionados (*error, fallo, incorrección*) toman como único sentido en un estudio de este tipo aquellas formas o usos que difieren del estándar en cuanto a los aspectos ortográfico, semántico, gramatical, combinatorio, diatópico, diastrático y diafásico. Dado que el análisis tiene como objeto de estudio la interlengua, no se deben inferir afirmaciones del tipo “los errores son malos” o “es un error, por lo que necesariamente debe ser corregido o erradicado en este nivel de dominio”.

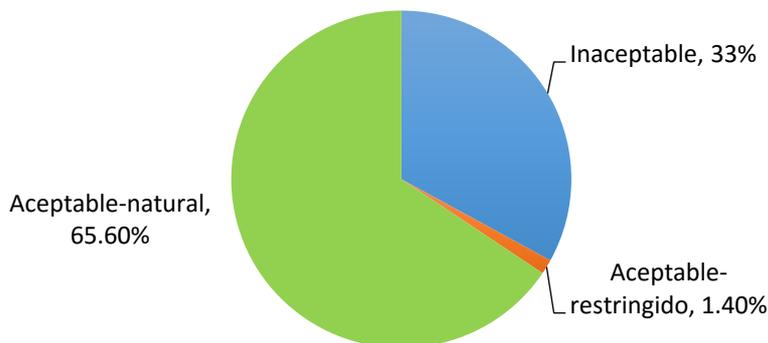


Gráfico 1. Índices de frecuencia de los usos aceptable-naturales, aceptable-restringidos e inaceptables

6.2. El error y el acierto en las estructuras analizadas

Aunque algunos de estos resultados no son significativos, pues las ocurrencias de ciertas estructuras son demasiado bajas como para que se puedan obtener resultados generalizables (13 son los errores en el núcleo verbal de las ditransitivas, pero 1 en las transitivas), resultan útiles como indicación a partir de la que comenzar una investigación más a fondo; por ejemplo, con una muestra de datos más amplia, como señalan McEnery y Hardie (2012, p. 51).

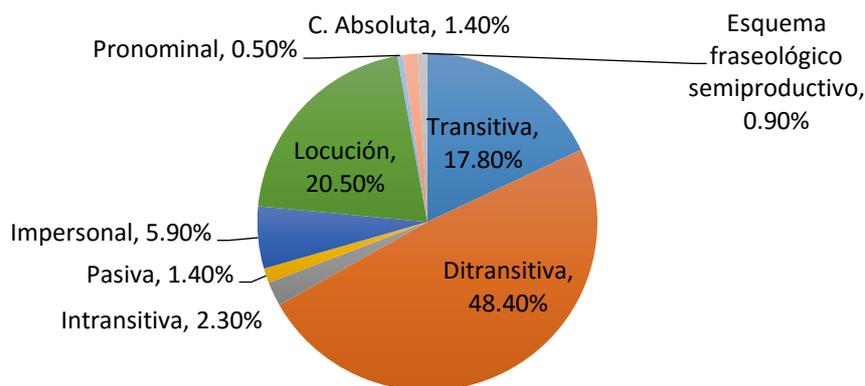


Gráfico 2. Porcentaje de error en el núcleo verbal de cada estructura

Así pues, estos resultados sí nos proporcionan algunos indicios. Las construcciones V + N son las que presentan el menor índice de errores (9 %), y su origen es fundamentalmente computacional¹⁰. Los errores de concordancia entre el verbo y el sujeto se dan también en otras estructuras, no es algo exclusivo de las ditransitivas; lo que sí es más propio de estas construcciones es el error en la restricción semántica del tema. De los errores computacionales vinculados al núcleo verbal, 5 se deben a la restricción semántica errónea; es decir, el 41,7 % de los errores relacionados con el núcleo del verbo en las oraciones ditransitivas es de esta naturaleza. Convendría contrastar estos resultados con un corpus más extenso. A continuación, y de acuerdo con la tipología del error

¹⁰ Entendemos por *error computacional* aquel que está relacionado con los rasgos de selección de la pieza léxica (problemas en la concordancia de rasgos); se refiere, por tanto, a cualquier problema relacionado con la selección sintáctica del verbo.

diseñada para clasificar los datos (basada en la de Corder, 1971) se incluyen los errores, su descripción y explicación:

- [1] Los judíos de España, llamados sefardíes, y los moros no nos daban el respeto que nos debieron
- [2] el museo fue establecido antes de su muerte y pues Miró le dió muchas obras. Otro día tomamos un tren a Montserrat)
- [3] Los viernes, siempre tuvimos cena, donde dimos bienvenidas a los nuevos estudiantes
- [4] Su éxito y la creación de una carrera que carece de sustancia da un mensaje negativo a otros
- [5] por la mañana hay camioncitos que vienen a dar cosas a las tiendas
- [6] Esta vez, uno de estos no funciona bien, y tenían que dar las cosas más tarde que regular
- [7] el pasar tiempo juntos nos daba mucha emoción aun cuando planeábamos
- [8] no sé que idiomas hablaré, en realidad no sé nada y eso me doy un poco miedo
- [9] Los pecados nos dan el muerte. !!!Jesus no da VIDA ETERNA!!!
- [10] es demasiado importante a dar algo a la comunidad que me has dado tanto en mi vida
- [11] un recuerdo que siempre tendré y siempre me dará una sonrisa
- [12] Cuando mi novio llegó del trabajo, le dió su regalo y nos preparamos
- [13] Emery y yo se dieron besitos cuando era medianoche porque es una tradición
- [14] y por eso tengo planes cuidarla bien, y darle todos los deseos de su corazón

En [3], se presenta una estructura ditransitiva formada por *dar*, con el sentido de “producir”, esto es, como verbo de apoyo, y un sustantivo que designa “recibimiento”, como es *bienvenida*. Este sustantivo pertenece a la clase de los sustantivos que designan “formas de trato social” (como los saludos, los recibimientos y las despedidas), junto con otros que se refieren a “diversos signos que forman parte de ceremonias”. Estos sintagmas nominales, cuando coaparecen con *dar*, están precedidos por un determinante (*dar el trato que se merece, dar el último adiós, darle la despedida, darle el recibimiento o la acogida merecida, el pésame, etcétera*), y el sustantivo aparece normalmente en singular –salvo interjecciones fosilizadas (*las gracias, los buenos días, las buenas tardes o las buenas noches*)–, dado que la acción se refiere genéricamente al tipo de recibimiento, excepto cuando se antepone una estructura comparativa, como en *la más calurosa de las bienvenidas o la peor de las bienvenidas*. Este error es intralingüístico y no puede atribuirse a la L1.

En [4] se demuestra que el aprendiente desconoce que los grupos nominales coordinados mediante la conjunción *y* que desempeñan la función de sujeto concuerdan con el verbo en plural: *Su éxito y la creación de una carrera que carece de sustancia da* (...). Este no es un caso en el que las dos ideas que se asocian se identifican en la conciencia lingüística del hablante como si designaran una sola entidad compleja, lo que justificaría el singular. El error es intralingüístico.

Los usos [8], [10], [12] y [13] se refieren a problemas en la concordancia de rasgos. En [8] y [12] parece darse el mismo patrón: el aprendiente flexiona el verbo en la persona y número del CI, en lugar de en los del sujeto: *Eso me doy un poco de miedo* y *[yo] le dio su regalo*. En [13], *Emery yo se dieron besitos*, el aprendiente confunde igualmente la 1.^a persona del plural con la 3.^a persona del plural. En [10] tampoco coinciden el sujeto (3.^a persona del singular) y el verbo (2.^a persona del singular). Estos errores, intralingüísticos, son ciertamente básicos, y se muestran en las redacciones aparentemente más simples desde una perspectiva lingüística y con más inadecuaciones.

Por último, destacan los errores en la restricción semántica del verbo, que se localizan en los usos [1], [7], [9], [11] y [14]. Con referencia a [1], en español, la palabra *respeto*, con el sentido de “veneración o consideración” (no con el sentido de “miedo o reparo”) no se da, sino que se muestra, se manifiesta, se profesa o se tiene. En cambio, en inglés sí se da el respeto, por lo que la fuente de este error es interlingüística, pues se está importando la restricción semántica. Lo mismo

sucede en los usos [11] y [14]. En español, la sonrisa no se da, pero en inglés sí (aunque tampoco se usaría en este contexto). En español sería más conveniente *provocar* para este contexto. Del mismo modo, tampoco se dan los deseos, sino que se conceden, excepto en la fórmula “le doy mis mejores deseos”. En inglés existe la expresión *to give a desire*, por lo que el aprendiente está importando esta restricción al uso el verbo español. En cambio, la fuente del error en [7] y [9] no proviene de la L1, por lo que el error es intralingüístico. La *emoción*, con el sentido de conmover el ánimo, no se da, sino que se provoca, se despierta, se causa. Por otro lado, con respecto al ejemplo [9] –*los pecados dan muerte*–, existe en español *dar muerte a alguien*, como *acabar con*; pero en este caso *muerte* nunca va precedido de determinante y el sujeto se tiene que poder personificar, lo que no es posible en el caso de *los pecados*. Además, se produce cierta incoherencia semántica, pues *dar muerte* se utiliza como sinónimo de *matar*, y los pecados no matan, sino que conducen a, causan o provocan la muerte.

Por otro lado, una menor cantidad de los errores son de origen léxico¹¹. En [5] y [6] el error léxico es de significado, pues consiste en una confusión en el significado de la pieza léxica seleccionada: la pieza léxica extraída existe pero tiene un significado diferente al que se le otorga. Así, el aprendiente confunde *dar* por *traer*; se trata de un caso de impropiedad semántica. El uso de *dar* en [2] es un error por el contexto comunicativo en el que se encuentra, debido a una desviación de la norma panhispánica en los factores extralingüísticos (geográficos, sociales y situacionales). Así, se produce un caso de pobreza léxica. La poca extensión de vocabulario del aprendiente causa la ocurrencia de este verbo frecuente en un contexto en el que un nativo –con el mismo perfil– no lo utilizaría; su grado de especialización semántica se halla por debajo de lo que un HN consideraría natural.

En suma, la interfaz léxico-sintaxis, a la que nos referíamos como posible foco de desviaciones, de acuerdo con las actuales hipótesis de ASL, es, en efecto, la principal fuente de errores en el uso transitivo y ditransitivo del verbo *dar* de este grupo de aprendientes de español de nivel avanzado, ya que el 78,6 % de los errores son computacionales, y no léxicos. Como se preveía, la desviación en la restricción semántica se convierte en el error computacional más frecuente (el 41,7 %), si bien los datos (14 errores de un total de 145 construcciones analizadas de HNN) son demasiado escasos para considerarlos significativos.

6.3. El papel de la transferencia de la L1

Por último, el análisis pretende responder a la última pregunta planteada en el primer apartado sobre el papel de la L1 en la producción escrita del *dar* transitivo y ditransitivo; no obstante, este análisis no nos permite asegurar si la influencia de la L1 tiene, en efecto, lugar, pues la influencia de la L2 es directa pero la de la L1 es indirecta. Teniendo esto en consideración y a juzgar por el análisis realizado, podríamos decir que el verbo *dar* en la construcción V+N no es, en general, muy susceptible de ser afectado por la transferencia de la L1 (de los 14 errores hallados en estas estructuras, solo 3 parecen deberse a la influencia de la L1), pero hay que matizar que la interferencia en la restricción semántica sí que puede ser muy importante, ya que la L1 está detrás del 60 % de los errores en la selección léxica del verbo; esto es, de los 5 errores debidos a un fallo en la selección léxica del verbo, 3 de ellos reflejan la influencia del inglés.

6.4. Conclusión del análisis

Para empezar, el verbo *dar* es un verbo infrautilizado por los HNN, como preveía Sinclair (1991) y como –a propósito del verbo *to make*– demuestran también los resultados de Altenberg y Granger (2001). No obstante, el estudio contrastivo de las estructuras sintácticas revela que los HNN sobreatilizan las ditransitivas, lo cual no resulta sorprendente en tanto que aparecen abundantemente en el input nativo, son las menos complejas formalmente y son sintácticamente

¹¹ A diferencia del error computacional, el error léxico se refiere a cualquier problema en la selección paradigmática. Se corresponde con dificultades en la adquisición de alguno de los aspectos del verbo: el semántico, el morfológico o confusión con su valor sociolingüístico.

equivalentes en la L1; además, caracterizan el léxico específico de la oralidad que es propio de los HNN en tareas escritas de tema libre.

Por otro lado, en las estructuras V + N el verbo actualiza el significado pleno de “proporcionar” y el de apoyo, “producir o hacer”. Cuando se utiliza el verbo como “proporcionar”, no hay diferencias muy llamativas entre ambos grupos; en cambio, sí las hay cuando el verbo actualiza el sentido de “producir(se)”. Se han distinguido dos usos del verbo de apoyo muy próximos: cuando se usa como “producir” en el sentido de “ejecutar” y cuando se usa como “producir” en el sentido de “suscitarse” (*Le dieron una patada* frente a *Le dieron miedo*). Los HNN sobreutilizan el sentido de “suscitarse”; en concreto, utilizan significativamente más el grupo de sustantivos que denotan “diversas sensaciones, percibidas como placenteras” y el grupo de sustantivos que denotan “miedo, inquietud y otros estados anímicos similares”. Estos resultados apoyan estudios que se refieren al mayor grado de implicación personal de los HNN en las producciones escritas cuando se comparan géneros equivalentes; esto explica que haya más referencias al estado mental y anímico del autor de la redacción. Llama la atención que, pese a esta sobreutilización, el conocimiento productivo extensivo de estas clases léxicas es menor, pues se usan menos sustantivos de la misma clase pero se repiten más; así, mientras los HN usan *miedo*, *reparo* o *pánico*, los HNN solo *miedo*; esto es aplicable a muchas de las clases léxicas: la variedad de sustantivos de la misma gama aporta una gran riqueza expresiva de la que carece el texto no nativo, y los HNN ganarían en naturalidad si los usaran. Por otro lado, los HN utilizan el verbo con una gran cantidad de sustantivos en sentido figurado, frente a los HNN, que apenas lo hacen (21,4 % frente a 1,6 %). La producción de sustantivos figurados es posiblemente más compleja, dado que la proyección metafórica que hace una comunidad de hablantes a partir de un sentido básico no tiene por qué coincidir con la de otra comunidad.

Por último, el análisis de errores y aciertos demuestra que no podemos dar por sentado que una construcción tan frecuente y aparentemente simple como es *dar* + N sea dominada por los HNN de nivel avanzado; que, si bien el contenido morfológico, semántico y sociolingüístico ocasionan pocos problemas, el gramatical (concordancia de rasgos) y sobre todo el lexicosintáctico (restricción semántica) pueden ser un importante foco de errores. La identificación de estas necesidades puede convertirse en punto de partida de aplicaciones en la enseñanza de español como lengua extranjera.

7. Conclusiones generales

Como se ha visto, analizar la lengua del aprendiente por medio de un análisis contrastivo de la interlengua produce resultados relevantes en la investigación sobre ASL. Los resultados obtenidos sugieren características estables en la interlengua de este grupo. No obstante, estos resultados deben ser, por un lado, triangulados en otros estudios experimentales o juicios metalingüísticos, ya que los corpus de datos constituyen muestras de actuación, no de competencia, por lo que solo permiten deducir cómo es la competencia de forma indirecta. Por otro lado, aquellos resultados no significativos por estar basados en escasas muestras de lengua deberían ser contrastados en corpus más amplios, pues el tamaño del CEDEL2, pese a ser un corpus bien diseñado, ha resultado en ocasiones insuficiente. Confiamos en que el tamaño de los corpus de aprendientes de español aumente en los próximos años. Sería útil también hacer un análisis de *dar* en otro corpus comparable de diferente L1 para confirmar que los rasgos que hemos identificado como interlingüísticos son verdaderamente influencias de la L1, y no consecuencia de patrones de desarrollo universales. Y, por último, convendría hacer un análisis manual del corpus que permitiera recuperar aquellos usos donde debería aparecer el verbo *dar*, pero en su lugar figura otro verbo; de esta manera se podría determinar si hay más casos de error en la restricción semántica de las CVA que en las de CVP, lo que constituía una de nuestras hipótesis de partida que no hemos podido contrastar. Estos cuatro estudios completarían esta investigación.

Referencias bibliográficas

- Altenberg, B. y Granger S. (2001). "The grammatical and lexical patterning of MAKE in native and non-native student writing". *Applied Linguistics*, 22 (2), 173-195.
- Bosque, I. (2001). "On the Weight of Light Predicates". En J. Herschenson *et al.* (Eds.), *Features and Interfaces in Romance. Essays in honor of Heles Contreras*, (pp. 23-38). Amsterdam: John Benjamins.
- Bosque, I. (Dir.). (2004). *Redes. Diccionario combinatorio del español contemporáneo*. Madrid: SM.
- Chi, M. *et al.* (1994). "Collocational problems amongst ESL learners: A corpus-based study." En L. Flowerdew y A. K. Tong (Eds.) *Entering Text* (pp. 157-165). Hong Kong: University of Science and Technology.
- Cobb, T. (2003). "Analyzing late interlanguage with learner corpora: Quebec replications of three European studies". *The Canadian Modern Language Review/La Revue canadienne des langues vivantes*, 59 (3), pp. 393-423.
- Corder, S. P. (1971). "Idiosyncratic dialects and error analysis". *International Review of Applied Linguistics*, 9, pp. 158-171.
- Davies, M. (2006). *A Frequency Dictionary of Spanish core vocabulary for learners*. Londres: Routledge.
- De Miguel, E. (2011). "En qué consiste ser verbo de apoyo". En M.V. Escandell Vidal *et al.* (Eds.), *60 problemas de gramática* (pp. 139-146). Madrid: Akal.
- Granger, S. (1996a). "Romance words in English: From history to pedagogy". *KVHAA Konferenser*, 36, pp. 105-121.
- Granger, S. (1996b). "From CA to CIA and back: An integrated contrastive approach to computerized bilingual and learner corpora". En K. Aijmer *et al.* (Eds.), *Languages in Contrast. Text-based crossed linguistic studies* (pp. 37-51). Lund: Lund University Press.
- Granger, S. (2015). "Contrastive Interlanguage Analysis: A reappraisal". *International Journal of Learner Corpus Research*, 1 (1), pp. 7-24.
- Gui, S. y Yang H. (2001). "Computer analysis of Chinese learner English". Conferencia pronunciada en el congreso Technology in Language Education, en Hong Kong University of Sciences and Technology. Recuperado de <<http://cle.ust.hk/common/conf2001/keynote/subject4/yang.pdf>> [consulta: 15-05-2016].
- Guo, X. (2006). *Verbs in the Written English of Chinese Learners: A Corpus-based Comparison between Non-native Speakers and Native Speakers* (Tesis doctoral). Universidad de Birmingham, Birmingham. Recuperado de <<http://etheses.bham.ac.uk/871/>> [consulta: 15-05-2016].
- Herrero Ingelmo, José L. (2002). "Los verbos soporte: el verbo *dar* en español". *Léxico y Gramática*. Recuperado de <<http://diarium.usal.es/joluin/files/2013/12/darsoporte.pdf>> [consulta: 15-05-2016].
- Hoffmann, S. y Lehmann H. M. (2000). "Collocational evidence from the British National Corpus." En J. Kirk (Ed.), *Corpora Galore. Analyses and Techniques in Describing English*. Amsterdam: Rodopi.
- Howarth, P. (1996). *Phraseology in English Academic Writing*. Tübingen: Max Niemeyer Verlag.
- Källvist, M. (1999). *Form-Class and Task-Type Effects in Learner English. A Study of Advanced Swedish Learners*. Lund: Lund University Press.
- Lapata, M. *et al.* (1999). "Determinants of Adjective-Noun Plausibility". En Proceedings of the 9th Conference of the European Chapter of the Association for Computational Linguistics, Bergen, 30-36. Recuperado de <<http://homepages.inf.ed.ac.uk/mlap/Papers/eacl99.pdf>> [15-05-2016].
- Lozano, C. (2009). "CEDEL2: Corpus Escrito del Español L2". En C. M. Bretones Callejas *et al.* (Eds.), *Applied Linguistics Now: Understanding Language and Mind* (pp. 197-212). Almería: Universidad de Almería.

- Lozano, C. y Mendikoetxea A. (2013). "Learner corpora and SLA: the design and collection of CEDEL2". En A. Díaz-Negrillo *et al.* (Eds.) *Automatic Treatment and Analysis of Learner Corpus Data* (pp. 65-100). Amsterdam: John Benjamins.
- Lozano Zahonero, M. (2011). *Gramática de perfeccionamiento de la lengua española*. Milán: Urlico Hoepli.
- Nagini, P. (2008). "The Mental Lexicon". *Journal of the Indian Academy of Applied Psychology*, 34 (1), pp. 181-186.
- Nation, P. (2001). *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nesselhauf, N. (2004). "Learner corpora and their potential for language teaching". En J. Sinclair (Ed.) *How to Use Corpora in Language Teaching* (125-152). Amsterdam: John Benjamins.
- Nesselhauf, N. (2005). *Collocations in a Learner Corpus*. Amsterdam: Benjamins.
- Ringbom, H. (1998). "Vocabulary frequencies in advanced learner English: A cross-linguistic approach". En S. Granger (Ed.), *Learner English on computer* (pp. 41-52). Londres: Longman.
- Ruhl, C. (1989). *On Monosemy: A Study on Linguistic Semantics*. Albany: State University of New York Press.
- Sánchez Rufat, A. (2010). "Apuntes sobre las combinaciones léxicas y el concepto de colocación". *Anuario de Estudios Filológicos*, 33, pp. 291-306.
- Sánchez Rufat, A. (2011). "Léxico gramaticalizado y lengua formulaica: algunas precisiones al enfoque léxico". *Sintagma*, 23, pp. 85-98.
- Sánchez Rufat, A. (2015a). "La naturaleza léxico-semántica del verbo *dar* en la construcción verbo + nombre". *Anuario de Estudios Filológicos*, 38, pp. 205-223.
- Sánchez Rufat, A. (2015b). El verbo *dar* en el español escrito de aprendientes de L1 inglés: estudio comparativo entre hablantes no nativos y hablantes nativos basado en corpus (Tesis doctoral). Universidad de Extremadura.
- Sánchez Rufat, A. y Jiménez Calderón, F. (2013). "Combinatoria léxica y corpus como input". *Language Design*, 14, pp. 61-81.
- Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Basingstoke: Palgrave.
- Schmitt, N. y Redwood, S. (2011). "Learner knowledge of phrasal verbs". En Meunier *et al.* (Eds.) *A Taste for Corpora. In honour of Sylviane Granger* (pp. 173-193). Amsterdam: John Benjamins.
- Shei, C. (1999). "A brief survey of English verb-noun collocation". Recuperado de <<http://www.dai.ed.ac.uk/homes/shei/survey.html>> [15-05-2016].
- Sinclair, J. M. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J. M. (2005). "How to build a corpus". En M. Wynne (Ed.), *Developing Linguistic Corpora: A Guide to Good Practice* (pp. 79-83). Oxford: Oxbow books.
- Viverg, A. (1996). "Cross-linguistic lexicology. The case of English *go* and Swedish *gå*". En K. Aijmer *et al.* (Eds.), *Languages in Contrast* (p. 155-182). Lund: Lund University Press.
- Wanner, L. *et al.* (2013b). "Annotation of Collocations in a Learner Corpus for Building a Learning Environment". En S. Granger *et al.* (Eds.), *Twenty Years of Learner Corpus Research: Looking back, Moving ahead*. Lovaina: Presses universitaires de Louvain. Recuperado de <http://www.dicesp.com/app/webroot/files/file/LCR2011_proceedings_wanner_leo_1_1.pdf> [15-05-2016].

Para citar este artículo

Sánchez Rufat, A. (2016). El uso transitivo y ditransitivo de dar en un corpus escrito contrastivo. *Revista Fuentes*, 18(2), 117-133. [Fecha de consulta: dd/mm/aa].
doi: <http://dx.doi.org/10.12795/revistafuentes.2016.18.2.01>