

# Complejidad y corrección de las construcciones con Ser/Estar en diferentes niveles de competencia del español como lengua extranjera

NURIA DE LA TORRE

*Instituto Cervantes de Estocolmo y Universidad Antonio de Nebrija (Madrid)*

MARIA CECILIA AINCIBURU

*Universidad Antonio de Nebrija y Università degli Studi di Siena*

Received: 26 November 2017 / Accepted: 01 April 2018

ISSN: 1697-7467

**Resumen:** Este artículo explora el uso de las construcciones con los verbos *ser* y *estar* para observar si existen datos empíricos que permitan asociar su aparición en producción no nativa con los niveles de competencia del Marco de referencia europeo. Se consideraron 124 tareas de escritura de una certificación académica realizada por estudiantes universitarios de intercambio. Los resultados muestran que hay características que marcan diferencias significativas entre los niveles, lo que se explica basándose en los conceptos de dificultad y complejidad del contexto. La consideración del uso correcto o incorrecto de los verbos *ser* y *estar* se asocia al contexto sintáctico e indica más errores en las combinaciones de *estar* que en las de *ser*. Los resultados generales muestran que algunas construcciones podrían usarse para distinguir niveles de competencia.

**Palabras clave:** *ser*, *estar*, niveles de competencia, evaluación, rasgos diferenciales, español como lengua extranjera

## Complexity and Accuracy with Ser/Estar in Different Levels of Competence in Spanish as a Foreign Language

**Abstract:** This paper explores the use of structures with the verbs *ser* and *estar* to see if there are empirical data that allow associating their appearance in non-native production, with the levels of linguistic competence proposed by the European Reference Framework. In the study, 124 writing assignments of an academic certification made by university exchange students were considered. The results show that there are features that mark significant differences between levels, for which explanations are advanced based on concepts of difficulty and complexity of the context. The consideration of correct or incorrect use of the verbs *ser* and *estar* in grammatical combinatorial is also associated. Candidates make more mistakes in the combinatorial of *estar* with respect to *ser*. The general results show that only some of the combinatorials could be used to distinguish levels of competence.

**Keywords:** *ser*, *estar*, levels of linguistic competence, evaluation, differential features, Spanish as foreign language.

## 1. INTRODUCCIÓN

El *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación* (2002 [2001], en adelante MCER) deja en manos de sus usuarios determinar “qué elementos, categorías, clases, estructuras, procesos y relaciones gramaticales tendrá que dominar el alumno”, ya que “no se considera posible elaborar una escala de la progresión relativa a la estructura gramatical que sea aplicable a todas las lenguas” (Consejo de Europa, 2002:111). Este hecho genera una importante línea de investigación puesto que, como explica Hulstijn (2010), la implementación correcta del MCER es prácticamente imposible si no se validan sus escalas, estableciendo vínculos claros entre las descripciones nociofuncionales y pragmáticas de los diferentes niveles y sus características lingüísticas. Este trabajo se inscribe en esa línea de determinar los correlatos formales de los niveles del MCER y explora si las estructuras que contienen los verbos *SER* y *ESTAR*, muy investigadas en el campo de la adquisición del español, pueden constituir un objeto de estudio válido para aportar datos relevantes a la descripción empírica de los niveles de referencia del MCER en español. Se analiza, por tanto, la frecuencia y distribución de esas construcciones y, cuando es posible, el grado de corrección en la selección de los verbos *SER* y *ESTAR* en 124 textos escritos por 62 participantes en el marco del pilotaje de un examen multinivel (B1-C1) de español académico de la Universidad Nebrija de Madrid, vinculado al MCER y validado por la Asociación ACLES.

## 2. LA COMPLEJIDAD Y LA CORRECCIÓN EN LA DESCRIPCIÓN DE NIVELES DE REFERENCIA

Los constructos de complejidad y corrección, estrechamente relacionados entre sí, se consideran descriptores de la actuación e indicadores del nivel de desarrollo en lenguas segundas (Bulté y Housen, 2012, Larsen-Freeman, 2009, Ortega 2012). La complejidad consiste en la habilidad que muestran los aprendices para usar un repertorio amplio y variado de estructuras y vocabulario de la lengua meta, así como en el carácter más o menos sofisticado de ese repertorio (Housen, Kuiken y Vedder, 2012; Ortega, 2015). Así Bulté y Housen (2012) consideran que factores como la dificultad relativa de procesamiento, la prominencia y frecuencia en el *input*, la relevancia comunicativa o la complejidad inherente provocan que un determinado rasgo lingüístico aparezca tardíamente, o no, en el desarrollo interlingüístico.

Por lo que se refiere a la corrección, Housen, Kuiken y Vedder (2012) la definen como el grado en el que la actuación de los aprendices se desvía de una norma dada. Aunque se trata de un concepto relativamente sencillo, su interpretación y aplicación a los datos de L2 puede resultar problemático. Además de las cuestiones derivadas de la gravedad relativa de los errores, puede ser complejo fijar los criterios por los que se considera que se producen desviaciones de la norma. Frecuentemente se considera norma el dialecto estándar del hablante nativo, pero en muchas ocasiones no existe una referencia estándar nativa o esta no es unánime en todas las comunidades y contextos sociales.

Aunque el crecimiento lineal de estos constructos a medida que se avanza de nivel no se puede dar por garantizado, ambos se utilizan para caracterizar los rasgos formales

de los niveles funcionales descritos en el MCER. Forsberg y Bartning (2010) analizan 83 muestras escritas en francés L2 por 42 estudiantes suecos, clasificadas por niveles mediante escalas basadas en descriptores del MCER. Los textos se sometieron al análisis y medición de rasgos morfosintácticos, discursivos y léxicos que ya se habían demostrado cruciales en estudios empíricos sobre los órdenes de adquisición del francés L2. El cálculo de errores señala diferencias significativas entre las producciones de los niveles A2, B1 y B2 y se observa un desarrollo lineal de ciertos rasgos discursivos y gramaticales.

Martin, Mustonen, Reiman y Seilonen (2010) hacen un estudio exploratorio sobre el finés L2 analizando 669 textos clasificados en los seis niveles del MCER. En su análisis documentan interacciones interesantes entre la emergencia, la frecuencia, la complejidad y la corrección en el uso de los recursos gramaticales.

Verspoor, Schmid y Xu (2012), desde una perspectiva basada en los sistemas dinámicos, analizan 437 textos escritos por aprendices holandeses de 12 a 14 años y muestran que los análisis lingüísticos de textos breves pueden proporcionar datos suficientes para distinguir entre los niveles inferiores de dominio (A1-B1) valorados por expertos con una escala holística derivada del MCER.

Mención aparte merecen los trabajos inscritos en el programa English Profile, un proyecto interdisciplinar que identifica rasgos lingüísticos capaces de discriminar la actuación de usuarios de los niveles del MCER a la luz de modelos psicolingüísticos de adquisición y utilizando avanzadas herramientas de lingüística de corpus. Salamoura y Saville (2010) resumen algunos hallazgos del programa en relación con los factores que determinan el orden de adquisición de esos rasgos, como son la frecuencia en la lengua meta, las jerarquías de complejidad semántica y estructural establecidas previamente en la lingüística del inglés o la L1 de los aprendices.

En el ámbito del español, de la Torre (2016), utilizando un corpus de 124 escritos académicos, analiza los textos en relación con mediciones de variedad verbal y encuentra correlaciones entre ellas y los niveles asignados a los textos por evaluadores expertos.

Los resultados de las investigaciones reseñadas, aunque casi nunca concluyentes, permiten justificar el rastreo de la aparición, frecuencia y corrección de determinadas estructuras en muestras de lengua clasificadas en diferentes niveles de competencia funcional para describirlos desde el punto de vista lingüístico. La elección de estudiar construcciones con *SER* y *ESTAR* se justifica porque se trata de uno de los aspectos más estudiados de la adquisición del español y por existir un relativo acuerdo entre los investigadores sobre los estadios en los que se desarrolla en la interlengua de aprendices de diferentes tipos. En el próximo apartado se condensa la investigación al respecto.

### 3. EL DESARROLLO DE LAS CONSTRUCCIONES CON *SER* Y *ESTAR*

La distinción entre los verbos *SER* y *ESTAR* se ha considerado tradicionalmente un aspecto problemático de la gramática del español. Una aproximación general indica que la selección de uno u otro verbo puede explicarse mediante criterios sintácticos, léxicos, semánticos y pragmáticos, o mediante una conjunción de varios de ellos. La descripción es particularmente compleja en el caso de las combinaciones con adjetivos, ya que algunos de los que son compatibles con *SER* también lo son con *ESTAR* bajo ciertas restricciones contextuales y discursivas.

En lo que se refiere al campo de la lingüística aplicada, la adquisición de los usos de *SER* y *ESTAR* ha constituido un fructífero objeto de estudio. Durante los años 80 y 90, los trabajos se inscribieron en la importante línea de investigación de la época dedicada a establecer órdenes y estadios transitorios de adquisición de diferentes fenómenos lingüísticos. Estudios como los de VanPatten (1985 y 1987), Ryan y Lafford (1992) y Guntermann (1992) van aportando datos para configurar unas etapas de desarrollo estables y predecibles que, a grandes rasgos, son las siguientes: (1) omisión del verbo y generalización de *SER*, (2) adquisición de *ESTAR* + gerundio, (3) aparición consistente de *ESTAR* + locativos y de *ESTAR* + adjetivos de estado y (4) emersión de la voz pasiva.

Con la entrada del nuevo milenio, gran parte de los estudios se centran en las construcciones con adjetivos y, al admitir que la mayoría de ellos se pueden combinar con ambos verbos, abandonan la categorización binaria y los juicios de corrección para dar cuenta de los procesos de adquisición. El objetivo de esta nueva línea de investigación es determinar los factores que mejor predicen la aparición de *ESTAR* + adjetivo en los distintos niveles de dominio (véanse, entre otros, Cheng, Lu y Giannakouros, 2008; Geeslin, 2003; Geeslin y Guijarro-Fuentes, 2004; Lu y Cheng, 2015).

Fuera del ámbito de las construcciones con adjetivos, de Garavito y Valenzuela (2008) usan juicios de gramaticalidad y selección de oraciones con 21 aprendices de nivel avanzado y 21 hablantes nativos para proponer la conclusión de que, debido a la carga de procesamiento que implican, las pasivas eventivas (*SER* + participio) se adquieren después que las resultativas y que las combinaciones con adjetivos.

Por último, Rodríguez López (2017) estudia la adquisición de las construcciones con *SER* y *ESTAR* en textos escritos por 239 aprendices alemanes con diferentes niveles de competencia (A1-C1). Este estudio tiene el mérito de utilizar las combinatorias de uso como resultan descritas en el Plan curricular del Instituto Cervantes (Instituto Cervantes, 2006; en adelante, PCIC), pero las pruebas analizadas por universitarios no implican el uso del Español académico. En el análisis se marcan errores/aciertos en la selección del verbo y la naturaleza sintáctica del predicado. En sus conclusiones Rodríguez López (2017: 232) indica que el índice de error en la selección del verbo es muy bajo (7%), pero que algunas estructuras, en concreto las combinaciones de *SER* + participio tienden a fosilizarse.

El PCIC propone una progresión en las combinatorias de *SER* y *ESTAR* con diferentes categorías gramaticales en todos los niveles de competencia, frecuentemente con una sección dedicada (Nociones gramaticales 12.1, Verbos predicativos). Como se espera en un texto de referencia, los usos no están avalados por estudios empíricos, pero presumiblemente coinciden con el orden en el que se enseñan institucionalmente (con detalle en Rodríguez López, 2017: 229).

#### 4. APROXIMACIÓN EMPÍRICA

De acuerdo con los objetivos del estudio, se formularon las siguientes preguntas de investigación:

- (1) ¿Existe diferencia en las estructuras que contienen *SER/ESTAR* cuando se usan en escritos de diferente nivel de competencia?
- (2) A medida que se avanza de nivel, ¿hay aumento significativo de la corrección en la selección de *SER/ESTAR* en contextos donde la opción por uno de los verbos es obligatoria?

En lo que se refiere a la primera pregunta, los tiempos de adquisición de las diferentes construcciones posibles con los verbos *SER/ESTAR* pueden tener alguna posibilidad de diferenciar los distintos niveles de dominio. Respecto a la segunda pregunta de investigación, la bibliografía disponible permite suponer que, incluso en los niveles más bajos, la selección de la cópula en contextos obligatorios muestra altos grados de corrección.

Para este trabajo se han analizado 124 textos de un examen multinivel (B1-C1) de español con fines académicos, escritos por 62 hablantes no nativos de español de diferentes niveles, con diferentes lenguas maternas<sup>1</sup>, que estudiaban temporalmente en la Universidad Nebrija de Madrid. Los textos proceden de la prueba de expresión e interacción escrita del examen, (60 minutos). Se realizaron manualmente dos tareas de escritura que luego se evaluaron aplicando escalas de calificación basadas en descriptores B1, B2 y C1 del MCER relacionados con el uso de la lengua en el ámbito académico. Se trata, por tanto, de un estudio explorativo, dados los escasos antecedentes en la investigación del Español Académico, y descriptivo. El hecho de que la evaluación certificativa forme parte de las actividades de formación universitaria aumenta el carácter ecológico de la investigación.

La tarea 1 (T1) consistía en escribir un correo electrónico a un profesor expresando desacuerdo con una nota recibida y formulando una solicitud oportuna. Se requería escribir un mínimo de 125 palabras. En la tarea 2 (T2) había que describir y explicar una tabla con datos estadísticos sobre el emplazamiento segregado de los alumnos con necesidades especiales en distintos países de Europa. Después el estudiante debía desarrollar un texto (entre 200 y 250 palabras) argumentando su postura personal. Entre las tareas hay diferencias claras en cuanto a extensión, tipología textual y requerimientos funcionales y pragmáticos, por lo que se mantiene por separado el análisis de los textos derivados de una y otra.

Un grupo de calificadoros expertos evaluaron los textos utilizando dos escalas de calificación de tipo Likert, una holística y otra analítica compuesta por cuatro categorías (alcance y control léxico y ortográfico, alcance y control gramatical, cohesión/coherencia discursiva y adecuación al género). El estudio estadístico de fiabilidad muestra un índice de coincidencia alto ( $\omega$  de McDonald = 0,870, asociado a una varianza de error inferior al 20%). Las escalas contenían cuatro bandas: la banda 1 describía la actuación escrita que no alcanzaba el nivel B1 del MCER, la 2 se destinaba a describir la producción de nivel B1, la 3 el B2 y la 4 el C1. Utilizando esas escalas, los textos quedaron clasificados en los tres niveles del MCER mencionados o por debajo de ellos, es decir, en el nivel A.

Para este trabajo se utilizó la clasificación de textos que resultó de la escala holística, que valora principalmente el logro del fin comunicativo de la tarea<sup>2</sup>. La distribución de los textos por niveles y tareas fue como sigue:

---

<sup>1</sup> Por la lectura de los textos se puede inferir que se trata de personas jóvenes de origen tan diferente como EEUU, China, Alemania o Inglaterra.

<sup>2</sup> Las calificaciones de los textos obtenidas en la escala holística y en las categorías de la analítica correlacionan fuertemente entre sí ( $p < .001$  en todos los casos).

Tabla 1. Distribución de los textos por niveles y tareas

	Nivel A	Nivel B1	Nivel B2	Nivel C1 <sup>3</sup>	
Tarea 1	10	25	22	5	n = 62
Tarea 2	12	24	25	1	n = 62
Total	22	49	47	6	n = 124

El número total de palabras de todo el corpus es de 23.496, 9.202 en la T1 y 14.294 en la T2.

Los textos se digitalizaron y se codificaron utilizando Atlas.ti. La codificación incluyó las siguientes categorías:

- (1) El contexto sintáctico al que el verbo vincula el referente. Se consideraron los siguientes:
  - (a) Contextos nominales (CN). Se codifican en esta categoría los sintagmas nominales, los pronombres, las subordinadas sustantivas y los numerales.
  - (b) Sintagmas adjetivales (SAdj.)
  - (c) Sintagmas adverbiales (SAdv.)
  - (d) Sintagmas preposicionales (SPrep.)
  - (e) Gerundios
  - (f) Participios
  
- (2) La aparición de error en la selección del verbo. Dado que el error en la selección de la cópula es opinable en determinados contextos sintácticos, se excluyen de este análisis los casos en los que la selección no es obligatoria. Se etiqueta *error* solo en los contextos nominales, donde es obligatorio SER, y en los siguientes contextos donde es obligatorio ESTAR: en la combinación con gerundio, en los contextos de sintagmas adverbiales y preposicionales cuando cumplen la función de localización espacial de referentes que no constituyen eventos y en la construcción con adjetivos de estado.

Se codificaron 800 casos de aparición de los verbos SER/ESTAR en cualquiera de sus formas, de los cuales se excluyeron, tal y como procede la investigación previa (ver Ryan y Lafford, 1992:716), 109 casos de instancias presumiblemente memorizadas que podían dificultar el rastreo de los procesos de adquisición: 9 fórmulas discursivas como “es decir” u “o sea”, 62 casos de “estar de acuerdo” y “estar a favor” que aparecían en las instrucciones de las tareas, 32 casos de la combinación “es posible” y 6 casos de fórmulas de cortesía del tipo “¿cómo estás?”. Los casos finalmente analizados fueron 691, 234 en los textos de la T1 y 457 en los de la T2.

<sup>3</sup> Dada la escasez de textos de nivel C1 se decidió excluirlo del análisis.

#### 4.1. La distribución de los predicados de SER/ESTAR

Como ya se ha indicado, las apariciones de SER/ESTAR se etiquetaron según los contextos sintácticos en los que aparecían. La tabla 2 muestra la distribución de su frecuencia en los textos según tarea y nivel.

Tabla 2. N° de casos y promedio (desviación típica) de construcciones con SER/ESTAR

ESTAR + gerundio (n=18)	T1 (n=13)	Nivel A (n=1)	Nivel B1 (n=10)	Nivel B2 (n=2)
	Media (DT)	0,01 (0,04)	0,08 (0,14)	0,03 (0,10)
	T2 (n=5)	Nivel A (n=1)	Nivel B1 (n=1)	Nivel B2 (n=3)
	Media (DT)	0,01 (0,02)	0,01 (0,03)	0,01 (0,03)
SER/ESTAR + sintagmas adverbiales (n=25)	T1 (n=10)	Nivel A (n=2)	Nivel B1 (n=6)	Nivel B2 (n=2)
	Media (DT)	0,04 (0,08)	0,05 (0,11)	0,01 (0,04)
	T2 (n=15)	Nivel A (n=7)	Nivel B1 (n=5)	Nivel B2 (n=3)
	Media (DT)	0,12 (0,18)	0,02 (0,05)	0,01 (0,04)
SER/ESTAR + sintagmas preposicionales (n=40)	T1 (n=7)	Nivel A (n=0)	Nivel B1 (n=3)	Nivel B2 (n=4)
	Media (DT)		0,02 (0,05)	0,04 (0,08)
	T2 (n=33)	Nivel A (n=4)	Nivel B1 (n=17)	Nivel B2 (n=12)
	Media (DT)	0,04 (0,09)	0,09 (0,14)	0,06 (0,09)
SER/ESTAR + participio (n=50)	T1 (n=12)	Nivel A (n=1)	Nivel B1 (n=5)	Nivel B2 (n=6)
	Media (DT)	0,02 (0,07)	0,03 (0,10)	0,05 (0,09)
	T2 (n=38)	Nivel A (n=3)	Nivel B1 (n=6)	Nivel B2 (n=29)
	Media (DT)	0,03 (0,07)	0,03 (0,09)	0,12 (0,15)
SER + contextos nominales (n=252)	T1 (n=90)	Nivel A (n=24)	Nivel B1 (n=31)	Nivel B2 (n=35)
	Media (DT)	0,61 (0,29)	0,27 (0,27)	0,47 (0,29)
	T2 (n=162)	Nivel A (n=25)	Nivel B1 (n=63)	Nivel B2 (n=74)
	Media (DT)	0,31 (0,28)	0,36 (0,23)	0,41 (0,27)
SER/ESTAR + sintagmas adjetivales (n=306)	T1 (n=102)	Nivel A (n=15)	Nivel B1 (n=49)	Nivel B2 (n=38)
	Media (DT)	0,32 (0,25)	0,57 (0,36)	0,39 (0,26)
	T2 (n=204)	Nivel A (n=38)	Nivel B1 (n=88)	Nivel B2 (n=78)
	Media (DT)	0,50 (0,36)	0,49 (0,26)	0,36 (0,23)

#### 4.1.1. *ESTAR + gerundio*

Aunque las tareas no parecen propiciar el uso de esta estructura (suponen el 2,60% de todos los casos), la tabla 2 muestra que el nivel B1, y en cierta medida el B2, recurren a ella en más ocasiones. En concreto, el B1 aporta más del 60% de los casos. En los textos de la T1 al menos la mitad de los casos que aparecen en B1 y todos los casos de A y B2 se pueden interpretar como una hipergeneralización de la perífrasis, que cubre funciones dentro de la redacción del mensaje que en español suele reservarse para otras formas verbales. Así, por ejemplo, se utiliza para iniciarlo (*Estoy escribiendo porque...*) o para concluirlo (*Estoy esperando su respuesta*). Esta observación coincide con el abuso de la estructura que documenta Guntermann (1992:1301) y que la investigadora explica por transferencia del inglés. Además, la práctica falta de casos en el nivel A, puede indicar que esta estructura se adquiere después de ese nivel.

#### 4.1.2. *SER/ESTAR + sintagmas adverbiales*

Suponen el 3,62% de los casos. Según los datos de la tabla 2, no se trata de una estructura especialmente problemática para nuestros aprendices ya que se documentan apariciones relativamente abundantes en los niveles inferiores.

#### 4.1.3. *SER/ESTAR + sintagmas preposicionales*

Constituyen el 5,79% de los casos. En la tabla 2 se observa su baja incidencia en el nivel A (solo cuatro casos) y la frecuencia relativamente más estable en B1 y B2 (20 y 16 casos respectivamente). Es posible que en el nivel A haya una tendencia a evadir estructuras con preposición, combinadas o no con estos verbos, y esto se refleje en nuestros datos.

#### 4.1.4. *SER/ESTAR + participio*

La tabla 2 muestra que la aparición de esta construcción es más relevante en la T2 (8,32% del total de casos) que en la T1 (5,13%) y que su frecuencia aumenta a medida que se avanza de nivel, siendo muy esporádica en el nivel A y llegando a tener una frecuencia relativamente importante en B2, sobre todo en los textos de la T2.

La presunta dificultad de la estructura<sup>4</sup> podría explicar su baja frecuencia en los niveles A y B1 y su crecimiento en el B2. De hecho, dos casos de los cuatro que se documentan en el nivel A, ambos de la misma informante, se podrían interpretar como intentos fallidos de construir pasivas verbales cuando todavía no se ha adquirido la estructura: *\*los alumnos con discapacidad son ayudando en sus estudios (...)* *\*los cuartos son arreglar por los estudiantes*. Gerundio e infinitivo intentan cumplir las funciones de un participio. En los niveles B1 y B2 se observa cómo va creciendo el uso de *SER* y, en prácticamente ningún caso, es posible discernir si se ha producido la selección correcta del verbo ya que se desconoce si la intención de los autores era la de expresar significados eventivos o resultativos<sup>5</sup>.

<sup>4</sup> Ver de Garavito y Valenzuela (2008) y el subconstructo de complejidad formal en Bulté y Housen (2012).

<sup>5</sup> Compárese, por ejemplo, *los alumnos con algún tipo de discapacidad estaban integrados con los alumnos sin discapacidad* con *\*no creo que todos los alumnos con NEE deba ser integrados con los emplazamientos segregados*. Ambos fragmentos corresponden a textos de sujetos diferentes clasificados en el nivel B2.

Lo que sí que parece que se está en condiciones de observar, al igual que se hizo con las perífrasis de gerundio, es que se produce una hipergeneralización de la estructura para cubrir funciones que en el habla nativa se implementan mediante la pasiva refleja u otro tipo de estructuras. La transferencia del inglés, como L1 o L2, puede ser la causa de este abuso. Algunos ejemplos son: *\*Eso debe solamente ser prohibido si los estudiantes (...)*<sup>6</sup> (en B1) o *\*El argumento por integración total no es desarrollado* (en B2). De hecho, se observa que los participantes en nuestro estudio probablemente transfieren del inglés propiedades de las construcciones pasivas que están vedadas en español. Mientras que la transformación pasiva en español está limitada a los verbos transitivos, en otras lenguas, como por ejemplo el inglés, la regla sintáctica de esta transformación es “ciega” y aparece en más contextos. Este es uno de los ejemplos que propone Demonte (1983, en Barahona 2013:18) cuando discute el carácter léxico o sintáctico de las pasivas en español:

- (a) *Someone gave a prize to John.* (Alguien dio un premio a John.)
- (b) *A prize was given to John.* (Un premio fue dado a John.)
- (c) *John was given a prize.* (\*John fue dado un premio)

La posibilidad de (c) en inglés es lo que quizá produce los siguientes ejemplos de nuestro corpus: *\*Si estos estudiantes no están permitido estudiar en estos emplazamientos (...)* (en B1) o *\*solo he perdido dos días de clase que estoy permitido* (en B2).

Los datos de los textos de la T2 indican que en los niveles A y B1 la estructura es esporádica y depende de unos pocos sujetos. Sin embargo, en B2 se generaliza a más participantes y la frecuencia aumenta ostensiblemente. El coeficiente de correlación de Kendall<sup>7</sup> es de 0,369, no muy fuerte pero estadísticamente significativo ( $p=0,001$ ), lo que evidencia desarrollo de la estructura a medida que se avanza de nivel. El contraste de Kruskal-Wallis muestra que existen diferencias entre los grupos ( $\chi^2 = (2, n=61)12,6, p = 0,002, \eta^2 = 0,136$ ) y las pruebas *post-hoc* (Dwass-Steel-Critchlow-Fligner) localizan la diferencia entre los niveles A y B2 y B1 y B2. Parece, por tanto, que el uso de la estructura *SER/ESTAR + participio* podría funcionar como indicador de haber alcanzado el nivel B2 del MCER cuando con la tarea propuesta se obtienen textos que propician su aparición.

#### 4.1.5. *SER + contextos nominales*

Se aborda ahora el segundo contexto más frecuente en nuestro corpus ya que aporta el 36,47% de los casos. En la tabla 2 llama la atención que en los textos de la T1 la estructura tiene una altísima frecuencia en el nivel A y una mucho más baja en B1. De hecho, las pruebas de varianza indican que hay diferencia entre ambos grupos (prueba de Kruskal-Wallis:  $\chi^2 = (2, n=56) 8,61, p = 0,013, \eta^2 = 0,178$ ).

---

<sup>6</sup> Se marcan con asterisco los ejemplos del corpus que contienen elementos que resultan idiosincrásicos en su contexto, aunque no siempre se pueda afirmar que se trate de secuencias agramaticales.

<sup>7</sup> Se utilizan en todo el trabajo pruebas estadísticas no paramétricas ya que los tests de normalidad (Shapiro-Wilk) arrojan resultados negativos. El análisis estadístico se ha llevado a cabo con el programa Jamovi (versión 0.9) disponible en <https://www.jamovi.org>.

#### 4.1.6. SER/ESTAR + sintagmas adjetivales

Estas construcciones son las más frecuentes de nuestro corpus (44,28% de los casos), y también las más estudiadas por la bibliografía. Su mera frecuencia de uso, sin atender a la selección de uno u otro verbo, no produce ni correlaciones con las calificaciones ni diferencias entre grupos.

#### 4.2. Los grados de corrección

Como se ha indicado, se mide la corrección solo en los contextos en los que hay selección obligatoria de uno de los verbos: SER + contextos nominales y las de ESTAR + gerundio, ESTAR + locativo (sin eventos como referentes, caso que no se encuentra entre nuestros datos) y ESTAR + adjetivo de estado.

Aunque en las combinaciones de SER + contextos nominales la precisión no sea total, la hipergeneralización de ESTAR, que en este contexto equivale al índice de error, es anecdótica:

Tabla 3. Porcentaje de casos de CN combinados con ESTAR

	NIVEL A	NIVEL B1	NIVEL B2	TOTAL
T1	8,33%	3,23%	0,00%	3,33%
T2	12,00%	6,35%	1,35%	4,94%
TOTAL	10,20%	5,32%	0,92%	4,37%

En la tabla 3 se observa en los totales que la corrección es solo ligeramente inferior al 90% en los textos de nivel A<sup>8</sup>. A medida que se avanza de nivel va retrocediendo hasta desaparecer casi por completo en el nivel B2. Es, por tanto, una estructura de temprana adquisición, como cabría esperar en una combinación que obligatoriamente requiere la forma no marcada. Así todo, la aparición del error revela una correlación inversa con el avance de nivel (coeficiente de Kendall -0,234 significativo en el nivel  $p=0,013$ ).

Por lo que respecta a los contextos de uso obligatorio de ESTAR, en las combinaciones con gerundio se encuentra un 100% de precisión en la selección del verbo, lo que coincide con los altos grados de corrección en esa estructura hallados en otros estudios (Rodríguez-López, 2017; Ryan y Lafford, 1992).

En lo que se refiere a los locativos, se hallan 26 casos en todo el corpus, 4 en combinaciones con adverbios (3 en B1 y 1 en B2) y 22 en combinaciones de sintagmas preposicionales (2 en A, 15 en B1 y 5 en B2). Todos ellos se combinan correctamente con ESTAR, excepto en dos casos de SPrep. de un mismo texto clasificado en el nivel B1 (en ese texto aparece otro caso de locativo acompañado de ESTAR). El porcentaje de acierto es de 92,3%, lo que puede indicar que es una estructura que se adquiere relativamente pronto, aunque, dada la escasez de los datos, puede haber estrategias de omisión que estén oscureciendo el resultado.

<sup>8</sup> 90% de corrección es el umbral que establecen VanPatten (1985, 1987), Ryan y Lafford (1992) y Guntermann (1992) para determinar que una estructura está adquirida.

Por último, los datos sobre ESTAR + adjetivo de estado son también muy escasos: hay 8 adjetivos de ese tipo, que dan lugar a 11 casos (3 en A, 5 en B1 y 3 en B2). Solo en dos de ellos se produce la generalización errónea de SER: *\*Soy disponible después de las 14:00* (en un texto del nivel B2 donde también aparece *\*Me gustaría hablar contigo (...) cuando estás disponible*) y *\*si el cuerpo sea completo* de un texto en el nivel B1. El porcentaje de acierto, por tanto, es del 81,8%, el más bajo de todas las construcciones analizadas en este apartado.

## 5. CONCLUSIONES

Las preguntas de investigación de este trabajo cuestionaban el hecho de si el estudio del desarrollo de las construcciones con los verbos SER/ESTAR podía aportar datos significativos a la descripción de las características formales de los niveles de referencia del MCER.

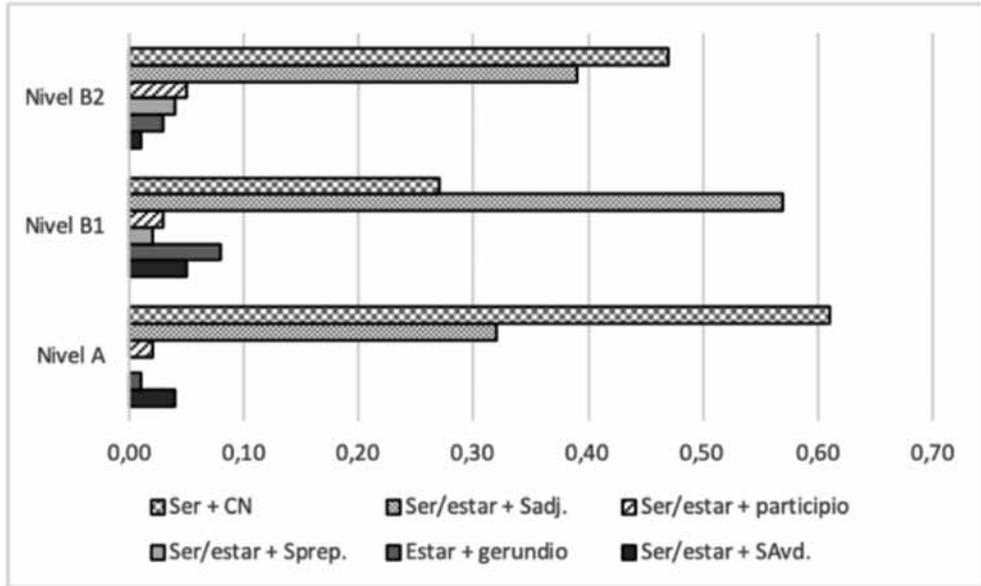
La primera hipótesis que se sostenía era la de que los tiempos de adquisición de algunas de las construcciones posibles con los verbos SER/ESTAR pudieran discriminar entre los distintos niveles de dominio. Y en este sentido se ha encontrado que la mayor frecuencia de la combinación SER/ESTAR + participio es un rasgo de los textos clasificados en el nivel B2, significativamente diferentes en este aspecto a los textos de nivel B1 y A. Este resultado está en consonancia con los previos de Guntermann (1992), de Garavito y Valenzuela (2008) y de Rodríguez López (2017). La complejidad inherente a la estructura y su baja frecuencia en el *input*, junto a una posible transferencia de L1 o L2, pueden explicar el resultado. De la misma manera, el PCIC señala las pasivas de acción (construidas con SER) como un contenido propio del nivel B2, mientras que las de resultado (construidas con ESTAR) lo serían de C1. Ha de tenerse en cuenta a este respecto que lo significativo en nuestros datos no es la precisión en la selección del verbo, imposible de juzgar en la mayoría de los casos sin conocer las intenciones de los autores, sino la aparición de la propia construcción, bien sea eventual o resultativa.

A propósito de la medición de la frecuencia de las diferentes construcciones, se ha encontrado un resultado inesperado y para el que resulta difícil hallar una explicación teórica. En el análisis de los textos de la T1, la combinación con contextos nominales (mayormente sintagmas nominales) mostraba una diferencia estadísticamente significativa entre los niveles A y B1. Una posible causa podría ser que en el nivel B1 empieza a aparecer una variedad de construcciones que no están presentes en el nivel A y que en el nivel B1 le roban protagonismo a la construcción con sintagmas nominales. En la figura 1 se observa que la frecuencia de combinaciones con contextos nominales es más baja en el B1, donde hay un protagonismo indiscutible de la combinación con adjetivos. También se puede ver que el resto de las combinaciones, menos las de los adverbios que decrecen a medida que aumenta el nivel, se incrementan en B1 respecto al nivel A. La falta de datos solo permite aventurar que en B1 se esté produciendo una suerte de reestructuración del sistema que hace que los aprendices experimenten con combinaciones más arriesgadas, como podrían ser los adjetivos o los sintagmas preposicionales, al mismo tiempo que dan cabida en su producción a construcciones recientemente adquiridas como las perífrasis con gerundio<sup>9</sup>.

---

<sup>9</sup> A propósito de las posibles peculiaridades del nivel B1, en los datos de Rodríguez López (2017) se observa que ese nivel supone una especie de punto de inflexión a partir del cual la frecuencia de errores, que hasta entonces crecía, comienza a decrecer linealmente.

Figura 1. Perfil de los niveles según la frecuencia de los contextos sintácticos en los que aparecen SER/ESTAR en los textos de la T1



Nuestra segunda hipótesis tenía que ver con los índices de selección correcta de los verbos en los contextos en los que uno u otro son obligatorios. La mayor parte de los datos sobre contextos obligatorios de nuestro análisis los aporta la construcción con sintagmas nominales, donde SER, la opción no marcada, es obligatorio. Partiendo de la base bibliográfica, se consideró que, en general, sería posible encontrar altos grados de corrección, aunque también sería posible observar cierta mejoría con el avance de nivel. Tras el análisis, la hipótesis se sostiene y está en consonancia con los resultados de Van Patten (1985 y 1987), Ryan y Lafford (1992), Guntermann (1992) y Lu y Cheng (2015) que observan altos grados de corrección en el uso del verbo SER. La corrección en ese contexto es superior al 90% en prácticamente todos los niveles, aunque se obtienen correlaciones significativas de aumento de la precisión a medida que se avanza de nivel. Los datos de contextos obligatorios de ESTAR son escasos en nuestro corpus y casi siempre son correctos. Dado el grado de corrección encontrado, y ya desde el punto de vista de la didáctica de estas construcciones, parece que es conveniente dirigir el foco, más que a la selección impropia del verbo, hacia la hipergeneralización de las construcciones ESTAR + gerundio y SER/ESTAR + participio, que cumplen funciones en los textos analizados que no realizan normalmente en el habla nativa. En este sentido, sobre todo en lo que se refiere a las combinaciones con participio, se observa que la mayor complejidad gramatical en sí misma no produce siempre textos de mayor calidad<sup>10</sup>.

<sup>10</sup> Ver Hyland (2002: 8–10, en Weir, 2005:116) sobre la falta de correspondencia entre la complejidad y la calidad de la escritura desde el punto de vista comunicativo.

A la vista de los resultados, y respecto al primer objetivo de este trabajo, es posible que el estudio de la selección de los verbos *SER/ESTAR* no aporte datos definitivos para describir los perfiles lingüísticos de los niveles de referencia. Sin embargo, algunas estructuras de las que forman parte estos verbos, como las combinaciones con participio, resultan significativas, por lo que debería ser posible continuar indagando en un corpus que contuviera los 6 niveles de competencia. Los resultados, de todos modos, deben tomarse con la debida cautela dadas las limitaciones del estudio. Como en trabajos anteriores, el análisis sugiere que las apariciones de *ESTAR* tienden más al error que las que contienen *SER*. Sería necesario, en ese sentido, un análisis mayormente cualitativo, con escrutinio de hablantes nativos, que incluyera las construcciones donde ambos verbos alternan y que permitiera conocer aspectos importantes del estado de la interlengua en los diferentes niveles.

## 6. REFERENCIAS

- Barahona Coroner, N. (2013). *La predicación con ser y estar en español*. (Unpublished master's thesis). Universidad Autónoma de Madrid, Madrid, Spain.
- Bulté, B. & Housen, A. (2012). Defining and operationalising L2 complexity, in A. Housen, F. Kuiken, & I. Vedder (Eds) *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA*. (pp. 21-44). Amsterdam: Benjamins.
- Cheng, A. C., Lu, H. C., & Giannakouros, P. (2008). The uses of Spanish copulas by Chinese-speaking learners in a free writing task. *Bilingualism: Language and Cognition*, 11(3), 301-317. doi: 10.1017/S1366728908003532
- Consejo de Europa (2002). *Marco común europeo de referencia para las lenguas: aprendizaje, enseñanza, evaluación*. Madrid: Grupo ANAYA. Available from: [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/marco/cvc\\_mer.pdf](https://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/cvc_mer.pdf), accessed 11 November, 2016.
- de Garavito, J.B. & Valenzuela, E. (2008). Eventive and stative passives in Spanish L2 acquisition: A matter of aspect. *Bilingualism: Language and Cognition*, 11(3): 323-336. doi: 10.1017/S1366728908003532
- de la Torre García, N. (2016). *La complejidad lingüística en los niveles de competencia del MCER: el caso de la variedad verbal en la expresión escrita en ELE*. *Revista Nebrija de Lingüística Aplicada*, 20(10): 4-21. Available from: <https://www.nebrija.com/revista-linguistica/la-complejidad-linguistica-en-los-niveles-de-competencia-del-mcer-el-caso-de-la-variedad-verbal-en-la-expresion-escrita-en-ele.html>, accessed 15 January, 2017. doi: 10.26378/rmlael1020260
- Forsberg, F. & Bartning, I. (2010). Can linguistic features discriminate between the communicative CEFR-levels?: A pilot study of written L2 French., in I. Bartning, M. Martin, & I. Vedder (Eds.) *Communicative proficiency and linguistic development: Intersections between SLA and language testing research*. Eurosla Monograph Series, 1. Available from: <http://eurosla.org/monographs/EM01/EM01home.html>, accessed 11 November, 2016.
- Geeslin, K.L. (2003). A Comparison of Copula Choice: Native Spanish Speakers and Advanced Learners. *Language Learning*, 53(4): 703-764. doi: 10.1046/j.1467-9922.2003.00240.x
- Geeslin, K.L & Guijarro-Fuentes, P. (2004). Estudio longitudinal del Ser y Estar en el español como L2. *Porta Linguarum*, 2, 93-110. Available from: <http://www.ugr.es/~portalin/>, accessed 6 Juni, 2017.
- Guntermann, G. (1992). An Analysis of Interlanguage Development over Time: Part II, ser and estar. *Hispania*, 75(5): 1294-1303.

- Housen, A., Kuiken, F. & Vedder, I. (2012). Complexity, accuracy and fluency: Definitions, measurement and research, in A. Housen, F. Kuiken & I. Vedder (Eds.) *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA* (pp. 1-20). Amsterdam: Benjamins.
- Hulstijn, J.H. (2010). Linking L2 proficiency to L2 acquisition: Opportunities and challenges of profiling research, in I. Bartning, M. Martin, & I. Vedder (Eds.) *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 233-238). Eurosla Monograph Series, 1. Available from: <http://eurosla.org/monographs/EM01/EM01home.html>, accessed 15 November, 2016.
- Instituto Cervantes. (2006). *Plan curricular del Instituto Cervantes: niveles de referencia para el español*. Instituto Cervantes. Madrid: Biblioteca Nueva. Available from: [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/plan\\_curricular/](https://cvc.cervantes.es/ensenanza/biblioteca_ele/plan_curricular/), accessed 6 Juni, 2017.
- Larsen-Freeman, D. (2009). Adjusting Expectations: The Study of Complexity, Accuracy, and Fluency in Second Language Acquisition. *Applied Linguistics*, 30(4): 579–589. doi: 10.1093/applin/amp043
- Lu, H. C. & Cheng, A. C. (2015). Estudio de verbos copulativos a partir de corpus de aprendices. *Porta Linguarum*, 23: 205–220. Available from: <http://www.ugr.es/~portalin/>, accessed 12 December, 2017
- Martin, M., Mustonen, S., Reiman, N. & Seilonen, M. (2010). On Becoming an Independent User, in I. Bartning, M. Martin, & I. Vedder (Eds.) *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 57-80). Eurosla Monograph Series, 1. Available from: <http://eurosla.org/monographs/EM01/EM01home.html>, accessed 11 November, 2016.
- Ortega, L. (2012). Interlanguage complexity, in B. Kortmann & B. Szmrecsanyi (Eds.) *Linguistic Complexity: Second Language Acquisition, Indigenization, Contact* (pp. 127–155). Munich: Walter de Gruyter.
- Ortega, L. (2015). *Syntactic complexity in L2 writing: Progress and expansion*. *Journal of Second Language Writing*, 29: 82–94. doi: 10.1016/j.jslw.2015.06.008
- Rodríguez López, P. (2017). Análisis de interlengua de los rasgos gramaticales y léxico-semánticos con los verbos ser y estar, in M.C. Ainciburu (Ed.) *La adquisición del sistema verbal del español. Datos empíricos del proceso de aprendizaje del español como lengua extranjera* (pp. 201–234). Frankfurt am Main: Peter Lang.
- Ryan, J. M. & Lafford, B. A. (1992). *Acquisition of lexical meaning in a study abroad environment: Ser and estar and the Granada experience*. *Hispania*, 75(3), 714-722.
- Salamoura, A. & Saville, N. (2010). Exemplifying the CEFR: criterial features of written learner English from the English Profile Programme, in I. Bartning, M. Martin & I. Vedder (Eds.) *Communicative proficiency and linguistic development: Intersections between SLA and language testing research* (pp. 101-132). Eurosla Monograph Series, 1. Available from: <http://eurosla.org/monographs/EM01/EM01home.html>, accessed 11 November, 2016.
- VanPatten, B. (1985). The Acquisition of ser and estar by Adult Learners of Spanish: A Preliminary Investigation of Transitional Stages of Competence. *Hispania*, 68(2): 399–406.
- VanPatten, B. (1987). The acquisition of ser and estar: Accounting for developmental patterns, in B. VanPatten, T. Dvorak & J. Lee, (Eds.) *Foreign language learning: A research perspective* (pp. 61-75). New York: Newbury House.
- Verspoor, M., Schmid, M. S. & Xu, X. (2012). *A dynamic usage based perspective on L2 writing*. *Journal of Second Language Writing*, 21(3), 239-263. doi: 10.1016/j.jslw.2012.03.007
- Weir, C.J. (2005). *Language Testing and Validation*. Hampshire: Palgrave Macmillan.