Revista Iberoamericana de Lingüística



nº 16 2021

R.I.L. 16

Fundador y Editor:
Ricardo de la Fuente Ballesteros (Universidad de Valladolid)

Co-editor:
Department of Modern Languages and Literatures (University of Texas at San Antonio)

Director:
Francisco Marcos Marín (University of Texas at San Antonio)

Comité asesor:
José Antonio Pascual (Real Academia Española)
Liliana Sánchez (Rutgers University)
José Camacho (Rutgers University)
Mariana Achugar (Carnegie Mellon University)
Brunello Natale di Cussatis (Università di Peruggia)
Luis Santos Río (Universidad de Salamanca)
Alfredo Torrejón (Auburn University)
Miguel Casas Gómez (Universidad de Cádiz)
† José Antonio Samper Padilla (Universidad de las Palmas de Gran Canaria)
Francisco Ocampo (University of Minnesota)
Francisco Javier Satorre Grau (Universidad de Valencia)
Antonio Salvador Plans (Universidad de Extremadura)
Mario Benvenuto (Università della Calabria)
Aurora Martínez Ezquerro (Universidad de la Rioja)

Comité de redacción:
María Jesús Leal (Hamline University)
Antonio Gragera (Texas State University, San Marcos)
Antonio Carrasco (Universidad de Valladolid)
Muhammad el-Madkouri Maataoui (Universidad Autónoma de Madrid)
Beatriz Soto Aranda (Universidad Rey Juan Carlos).
Nelsy Echávez-Solano (College of Saint Benedict / S. John's University)
Rossella Michienzi (Università della Calabria)
Beatriz Valverde Olmedo (Fu Jen Catholic University [Taiwán])

Miembro de Honor:
Humberto López Morales (Secretario de la Asociación de Academias de la Lengua Española)

Edición:
Universitas Castellae y Alimentelamente
Apartado 37, 47080 Valladolid
Tlfno. 34 637 665 048
E-mail: info@alimentelamente.es
www.alimentelamente.es
www.universitascastellae.es
www.reviblin.com
ISSN: 1887-407X
Impresión: Lamsa Soluciones Gráficas

ARTÍCULOS

EL ORDEN VS EN ESPAÑOL: EXPLICACIÓN PARAMÉTRICA PARA SINOHABLANTES

THE ORDER VS IN SPANISH: PARAMETRIC EXPLANATION FOR SPEAKERS OF CHINESE

Beatriz Valverde Olmedo
Fu Jen Catholic University (Taiwán)

Recibido: 17 de septiembre 2021. Aceptado: 1 de diciembre 2021

Resumen: el principio semántico de inacusatividad es universal, no así el comportamiento sintáctico de los predicados inacusativos los cuales exhiben disposiciones estructurales diferentes que varían de unas lenguas a otras. Esta diversidad paramétrica constituye el núcleo de este estudio, dado que en nuestro contexto de enseñanza confluye la siguiente terna idiomática: L_1-chino mandarín, L_2-inglés y L_3-español. En este último ciertos verbos inacusativos son proclives al orden Verbo-Sujeto, por su parte, el inglés posee unos requisitos estructurales más rígidos que impiden la posposición del sujeto en aquellos casos en que el español sí la permite. Este diferente comportamiento ordenativo de algunos inacusativos en estas dos lenguas es fundamental abordarlo en nuestras clases de ELE para sinohablantes, puesto que en este escenario los alumnos utilizan no el molde de su L_1, sino el de su L_2, el cual, como acabamos de mencionar, presenta un grado de flexibilidad más restringido y esto puede originar una manifiesta transferencia negativa en sus producciones a nivel estructural.

Palabras clave: enseñanza de ELE, inversión del sujeto, orden de palabras, sinohablantes, verbos inacusativos.

Abstract: the semantic principle of unaccusative is universal, but not the syntactic behavior of unaccusative predicates which ex-

hibit different structural dispositions that vary from one language to another. This parametric diversity constitutes the nucleus of this study, since in our teaching context the following idiomatic triad converges: L1-Mandarin Chinese, L2-English and L3-Spanish. In the latter, certain unaccusative verbs are prone to the Verb-Subject order, for its part, English has more rigid structural requirements that prevent the postponement of the subject in those cases in which Spanish does allow it. This different ordering behavior of some unaccusatives in these two languages is essential to address in our ELE classes for non-speakers, since in this scenario the students use not the template of their L1, but that of their L2, which, as we have just mentioned, presents a more restricted degree of flexibility and this can cause a manifest negative transfer in their productions at a structural level.

Keywords: Chinese-speakers, ELE teaching, subject inversion, unaccusative verbs, word order.

1. Introducción

La conocida taxonomía lexicosemántica de los predicados intransitivos divide a estos en dos macroconjuntos: el de los inacusativos y el de los inergativos, distinción sustentada en la relación verbo-argumento. En los primeros el argumento es equivalente al objeto directo de un predicado transitivo en un nivel subyacente (Perlmutter 1978, Burzio 1986, Levin y Rappaport-Hovav 1995, Mendikoetxea 1999a). Esta equiparación entre el único argumento de un predicado monovalente y el argumento-objeto de uno bivalente se manifiesta paramétricamente en la estructura superficial, esto es, si bien los verbos inacusativos, y por extensión el principio de inacusatividad, son universales, su plasmación sintáctica varía de unas lenguas a otras. Dicha variación constituye el foco de atención en este estudio, concretamente, centrándonos en las tres lenguas presentes en nuestro ámbito educativo: chino mandarín (L_1), inglés (L_2 y lengua puente) y español (L_3)[1], nuestra pretensión radica en cómo abordar en nuestras clases de ELE para

sinohablantes[2] el reflejo sintáctico que produce la inacusatividad en español, puesto que en esta lengua dicho reflejo a nivel ordenativo provoca que los predicados inacusativos sean tendentes[3] a mostrar un orden no marcado informativamente de Verbo-Sujeto (VS), el cual se contrapone al patrón básico[4] Sujeto-Verbo-Objeto[5] (SVO) con el que se cataloga a esta lengua dentro del marco de la tipología lingüística.

El interés de tratar este tema reside fundamentalmente en los beneficios que aporta a los estudiantes una adecuada comprensión del orden de constituyentes "[…] a knowledge of the uses of word order as a syntactic device may help the student to understand structural relationships and to gain skill in forming good sentences [además de producir estructuras semejantes a las de un nativo]" (Watts 328-329)[6], aun así, es palmario el ostracismo al que está sometida tanto la enseñanza del orden de palabras[7] en ELE en general como aspectos tan concretos como la repercusión que posee el rasgo semántico de la clase verbal sobre la disposición del sujeto, en palabras de Hertel:

> […] explicit classroom instruction on Spanish word order is insufficient for acquisition. The word order of declarative sentences is rarely, if at all, discussed in Spanish classes or textbooks. A review of top-selling beginning, intermediate and advanced textbooks revealed that very few textbooks mentioned inversion in declarative sentences, and when they did reference was made only to the flexible nature of Spanish word order. No textbooks discussed the specific effects of verb classes or discourse factors on word order. In addition, the majority of Spanish instructors are not aware of the unaccusative-unergative distinction and its structural effects. (284)

Asimismo, a lo mencionado en el párrafo precedente, se suma el hecho de que considerar la tendencia de ciertas es-

tructuras inacusativas hacia el orden VS adquiere especial relevancia en nuestra esfera educativa. Esta está constituida por estudiantes universitarios sinohablantes taiwaneses de ELE[8] en la que el inglés es la L_2 preponderante[9] que influye –las más de las veces en forma de transferencia negativa– sobre la adquisición de otras lenguas L_3, L_4, etc., si estas son, como el inglés, lenguas occidentales y más concretamente europeas, pues los alumnos tienen la inconsciente y a su vez atinada intuición de que su lengua materna es radicalmente opuesta de aquellas[10], lo que les induce a establecer una red de comparaciones entre las lenguas que les parecen más semejantes, llegando a considerarlas equiparables:

> Si tomamos el ejemplo del español y el chino, lenguas sin apenas conexión, podemos encontrarnos con que muchos estudiantes chinos pretenden aplicar estructuras morfológicas o gramaticales de otras lenguas occidentales que han estudiado previamente (especialmente, del inglés) a la hora de construir una frase en español, y les resulta extremadamente complicado entender que lo que en una lengua occidental es correcto no lo es en la otra […] (Méndez Marassa 24)

En definitiva, en este contexto la L_2-inglés tiene más repercusión sobre la L_3-español que la L_1-chino mandarín sobre esta última[11], siendo esta repercusión esencial. Usualmente los estudiantes de una L_2 dada tienden a utilizar la plantilla gramatical de su L_1 (Herschensohn 2000, Schwartz y Sprouse 1996), hecho que principalmente se produce en las primeras etapas de su proceso de adquisición o cuando a los alumnos se les expone a cierta fenomenología novedosa y nunca antes abordada en sus clases; no obstante, en este escenario educativo en el que nos hallamos, los alumnos usan y transfieren no desde el molde de su L_1, sino desde el de su L_2, el cual se caracteriza por exhibir un grado de flexibilidad a nivel de orden más restringido que el del español. En

lo referente a los predicados inacusativos y a la disposición VS, si bien el chino mandarín y el español tienen un comportamiento en algún punto coincidente, como veremos más adelante, no así el inglés en el que la inacusatividad sintácticamente no es manifiesta "Although the same verb classes [la clase de inacusativos] exist in English (i.e., they are universal), their word order does not vary because English has a relatively fixed SV order. Unlike Spanish, English displays few, if any, syntactic reflexes of unaccusativity" (Hertel 274). Esta rigidez puede provocar que los estudiantes transfieran negativamente la horma sintáctica SV del inglés al español, obviando la clase verbal, lo que conllevaría una inadecuada sobregeneralización; por consiguiente, el objetivo último de este estudio se centra en controlar las producciones estructurales de nuestros estudiantes dadas las peculiaridades de nuestra ámbito de enseñanza.

2. Verbos inacusativos y orden VS: una cuestión paramétrica

Las tres lenguas en las que se basa este estudio se identifican como lenguas de sujeto inicial, y tanto el inglés como el español presentan, según la clasificación greenbergiana (1963), el orden básico SVO. Más difícil es atribuir al chino mandarín uno de los patrones propuestos desde la perspectiva de la tipología lingüística[12], como señalan Li y Thompson:

> Mandarin is not an easy language to classify in terms of word order, for three reasons […] First, the notion of subject is not a structurally well-defined one in the grammar of Mandarin[13]. A second fact […] is that order in which basic words and phrases occur is governed to a large extent by considerations of meaning rather than of grammatical functions. This means that sentences with verbs at the beginning, in the middle, and

> at the end can be found in Mandarin. Languages that are relatively easy to characterize in Greenberg's terms are always those in which word order is determined primarily on strictly grammatical grounds [...] Third, whether it is taken to verb medial or verb final, Mandarin is inconsistent with respect to the features that correlate with VO or OV order according to Greenberg's typological scheme. (19)

El hecho de que esta lengua se considere SVO-SOV corrobora el parecer de ciertos autores críticos con respecto a la asignación de un único patrón básico a cada lengua, o directamente el poder atribuir uno, entre los que se encuentra Dik:

> Since order is not a deep property of natural languages, there is no reason for there to be only a single "basic" order for a given language. The FG [gramática funcional] approach to constituent ordering is fully compatible with the co-existence of different patterns or templates, to be used in different conditions and for different purposes. (394)

Al margen de estas cuestiones que superan los límites del presente trabajo, concentramos nuestra atención en los verbos inacusativos y su configuración sintáctica en el trío idiomático que particulariza nuestro entorno educativo.

En español los predicativos intransitivos muestran una variación ordenativa entre las disposiciones SV y VS (Delbecque 1991, López Meirama 1997b, Melis y Alfonso Vega 2013). Esta variabilidad se debe a las propias características consustanciales de tales predicados:

> [...] la intransitividad es un dominio estructural flexible para la movilidad del participante sujeto debido a que este no entra en colisión con otro participante central, como ocurre en la cláusula tran-

> sitiva, donde sujeto y objeto tienden a la máxima diferenciación, incluyendo la posición respecto al verbo. Aun así, también en esta hay sujetos posverbales, pero en menor medida. (Suárez 164)

De estas dos distribuciones, anteposición SV y posposición VS, los verbos inacusativos se inclinan hacia esta última, especialmente aquellos que desempeñan una función presentativa (Suñer 1982). Según Alonso-Cortés, esta propensión de ciertos inacusativos se relaciona directamente con el papel temático que cumple el sujeto:

> La hipótesis activa que proponemos sostiene que en español el orden $V_{INTR.}$ – FN [frase nominal] S sirve para para marcar que el sujeto FN S es un caso de macropapel semántico paciente, mientras que el orden FN S – $V_{INTR.}$ indica que FN S es un caso de macropapel actor […]. Por tanto, el orden S–V frente a V–S es funcionalmente significativo en español porque marca el diferente tipo de sujeto del verbo intransitivo. (192)

En otras palabras, este autor nos propone diferenciar "los sujetos con propiedades de actor (A., S_A) de los objetos y sujetos con propiedades de paciente (O, S_O)" (195), como vemos, una vez más, se vuelve a incidir en la idea de equiparar el sujeto de los verbos intransitivos con el objeto de los transitivos, asociación que en esencia equivale al planteamiento de Dixon sobre el concepto de ergatividad: "A language is said to show ergative characteristics if intransitive subject [abreviado con la letra S] is treated in the same manner as transitive object [abreviado con la letra O], and differently from transitive subject" (60-61).

Acabamos de exponer sucintamente esa tendencia de algunas estructuras inacusativas por el orden VS, pero debemos tener en cuenta que "[…] unaccusativity is a necessary but not a sufficient condition for VS production" (Lozano y

Mendikoetxea 482), esto es, el principio semántico de la inacusatividad no implica forzosamente la posposición del argumento-sujeto. Una afirmación de este calibre, además de falaz, implicaría que la posición del sujeto en español se supedita exclusivamente a la clase verbal y esto, como toda la fenomenología que engloba el orden de palabras, no admite tales reduccionismos, pues esta disciplina es una amalgama interactiva de factores de diversa índole, especialmente, morfosintácticos, semánticos y pragmáticos.

Otro rasgo semántico asociado a la posposición del sujeto tienen que ver con la definitud, concretamente con los llamados sujetos escuetos o desnudos carentes de determinación, y en ocasiones de modificación: "[...] si el sujeto no tiene determinante ni complementos ha de aparecer obligatoriamente en posición posverbal" (Fernández Soriano 125), de ahí que "Posición del sujeto y determinación no son rasgos aislados: en concreto, existe una estrecha relación entre sujeto indeterminado y posposición [...]" (Mendikoetxea, *Construcciones con* se 1658). Este fenómeno no es exclusivo del español, sino que se ha de entender como un principio universal "[…] that the definite internal argument NP [sintagma nominal] is incompatible with inherent Case (cf. Belletti 1988; Lasnik 1992), and no natural human language allows a definite internal argument of the unaccusative verb to remain in direct object position dominated by V'[14]" (Yuan 291).

Nos hemos referido a la indefinitud de los SSNN en función de sujeto, obviando aquellos en función de objeto, dado que "Una observación que conviene tener presente es que esta restricción de determinación afecta exclusivamente a los sujetos: los objetos sin determinante son perfectamente lícitos en español […] Veo niños, Tengo problemas. He comprado patatas" (Fernández Soriano 126). Esta indeterminación que exhiben algunos argumentos-objeto incide directamente sobre el hecho de que los predicados inacusativos acepten sujetos escuetos precisamente por esa equivalencia: argumento–sujeto predicado intransitivo ≡ argumento–objeto predicado transitivo. Por su parte, los verbos inergativos requieren de

sujetos determinados en posposición, salvo en los casos de "inversión locativa" (Mendikoetxea 1999a). Ejemplo inventado: mientras que *¿?juegan niños* sería un ejemplo dudoso y con un claro orden marcado informativamente, al agregarle un locativo[15] *en el parque/en el patio del colegio juegan niños* se pierde esa marcación informativa.

La manifestación sintáctica de inacusatividad en chino mandarín es similar a la que presenta el español:

> It has been proposed that in Chinese, unaccusative verbs are distinguishable from unergative verbs (Huang 1990; Li 1990; Yu 1995). And there is abundant evidence indicating that the single argument of an unaccusative verb in Chinese can remain in object position at PF [forma fonética] […] as well as in preverbal position, […] although in the former case, the argument NP [sintagma nominal] must be indefinite. In contrast, the single arguments of unergative verbs are always in preverbal position whether they are definite or indefinite NPs, and they are not allowed in postverbal position […] (Yuan 279)

Veámoslo más claramente en la siguiente tabla que hemos confeccionado a partir de ejemplos adaptados de Yuan (279-280):

SSNN SUJETO INDETERMINADO + VERBO INACUSATIVO	
orden VS	沉沒三艘船 Pred.: 沉沒 Suj.: 三艘船 (Pinyin: chénmò sān sāu chuán – Trad. *Se hundieron tres barcos*)

orden SV	三艘船沉沒 Suj.: 三艘船 Pred.: 沉沒 (Pinyin: sān sāu chuán chénmò – Trad. *Tres barcos se hundieron*)
SSNN SUJETO DETERMINADO + VERBO INACUSATIVO	
orden VS ¿?	¿? 沉了那艘船 Pred.: 沉了 Suj.: 那艘船 (Pinyin: chénle nà sāu chuán – Trad. *Se hundió ese barco*)
orden SV	那艘船沉了 Suj.: 那艘船 Pred.: 沉了 (Pinyin: nà sāu chuán chénle – Trad. *Ese barco se hundió*)

Tabla 1. Posición de los sujetos (in)determinados con verbos inacusativos en chino mandarín

El ejemplo de SN sujeto determinado pospuesto con verbo inacusativo solo sería válido en los casos de la mencionada inversión locativa como sucede en español, de tal modo que si incluimos un circunstancial como 在這個海域 (pinyin: zài zhège hǎiyù, trad. *En esa zona marítima*) el sujeto se ve forzosamente desplazado 在這個海域沉沒了那艘船 (pinyin: zài zhège hǎiyù chénmòle nà sāu chuán, trad. *En esa zona marítima se hundió ese barco*).

Sobre los verbos inergativos la situación es más simple, independientemente de si el SN sujeto es indeterminado o no, este ha de ocupar la posición preverbal: 幾個孩子跳 (pinyin: jǐ ge háizi tiào, trad. *Algunos niños saltan*) o 這些孩子跳 (pinyin: zhèxiē háizi tiào, trad. *Estos niños saltan*), siendo inviables las distribuciones ¿?跳幾個孩子 (trad. *Saltan algunos niños*) y ¿?跳這些孩子 (trad. *Saltan estos niños*).

Por último, sobre el comportamiento de los inacusativos, la rigidez de la sintaxis inglesa con respecto a las otras dos lenguas examinadas se manifiesta en el hecho de que la variabilidad SV-VS de ciertos inacusativos se diluye en esta

lengua: "In the argument structure, the internal argument is inside VP [sintagma verbal], but English syntax does not allow it to remain *in situ* in the syntactic structure; this internal argument obligatorily moves to the preverbal subject position of the sentence to be Casemarked" (Yuan 277), el objetivo de este movimiento es satisfacer el Principio de Proyección Ampliado (PPA)[16], en otras palabras, el inglés precisa que en las oraciones con verbos finitos los sujetos antepuestos al verbo sean expresos. Así los ejemplos que hemos adaptado de Yuan en inglés se reducen todos al orden SV: *three ships sank, that ship sank*, *a few children jump* y *these children jump*, respectivamente.

A los dos factores semánticos analizados, inacusatividad y sujetos escuetos que favorecen la posposición del sujeto, añadimos uno de corte formal que también propicia esta ordenación, y que se puede concebir como el fenómeno antitético del segundo: los constituyentes de complejidad incremental o pesados, los cuales propenden a desplazarse a la última posición, según Dik: "[…] *The Principle of Increasing Complexity*. There is a preference for ordering constituents in an order of increasing complexity […] Complexity in this sense is often referred to as «heaviness» (Mallinson-Blake 1981, Hawkins 1983)" (404). Si bien esta complejidad o pesantez se puede aplicar a cualquier constituyente a nivel clausular, en nuestro caso nos centramos únicamente en los SSNN sujeto complejos, los cuales, a diferencia de los sujetos escuetos que son proclives a acompañar a la clase verbal de los inacusativos, pueden aparecer con cualquier tipo verbal. Véase la siguiente tabla con una serie de ejemplos inventados:

Sujeto complejo + Verbo inacusativo	
orden VS	*Ha aparecido el perro de color marrón que se perdió hace cinco días en el bosque de mi ciudad.* 出現了那隻五天前在我城市的森林裡迷路的棕狗[17] Pred.: 出現了 Suj.: 那隻五天前在我城市的森林裡迷路的棕狗 (Pinyin: chūxiànle nà zhī wǔ tiān qián zài wǒ chéngshì de sēnlín lǐ mílù de zōng gǒu)
orden SV	*¿?El perro de color marrón que se perdió hace cinco días en el bosque de mi ciudad ha aparecido.*
Sujeto complejo + Verbo inergativo	
orden VS	*Ha hablado el representante de la asociación vecinal que fue elegido por mayoría absoluta en la votación de la semana pasada.* 發言了那個在上周投票中以絕對多數當選的居委會代表 Pred.: 發言了 Suj.: 發言了那個在上周投票中以絕對多數當選的居委會代表 (Pinyin: fāyánle nàgè zài shàng zhōu tóupiào zhōng yǐ juéduì duōshù dāngxuǎn de jūwěihuì dàibiǎo)
orden SV	*¿?El representante de la asociación vecinal que fue elegido por mayoría absoluta en la votación de la semana pasada ha hablado.*

Tabla 2. Disposiciones de los sujetos complejos en español y chino

La flexibilidad exhibida por el español y el chino en la tabla 2 se ve constreñida en el inglés, el cual, una vez más, dadas sus características, no admite la posposición de estos sujetos pesados, por lo que los ejemplos de esta tabla se reducen al orden SV: *The brown dog that was lost five days ago in the forest of my city has appeared* y *The representative of the neighborhood association who was elected by an absolute majority in last week's vote has spoken*.

3. Orden VS: aplicación metodológica en el aula de ELE

En la sección anterior hemos expuesto tres fenómenos que producen la disposición VS en español: predicados inacusativos, SSNN sujetos escuetos y complejos, a partir de los cuales surge la siguiente pregunta: ¿cómo podemos realizar una aplicación metodológica en nuestra aula de español de este orden en función de los fenómenos contemplados? Uno de los principales objetivos de esta aplicación es que los alumnos sean conscientes de la flexibilidad en el orden de los constituyentes de unas lenguas a otras, y en su caso concreto del español, el chino mandarín y el inglés; asimismo se debe limitar la transferencia directa del rígido orden SV del inglés y que, por consiguiente, no produzcan una sobregeneralización, es decir, ni que apliquen el orden SV a todos los predicados monovalentes, ni lo contrario como le sucedió a Yuan en su estudio: "However, the data from groups 2–3 suggest that these groups of learners overgeneralized the use of postverbal arguments of the unaccusative; they incorrectly allowed the argument of the unergative verb and the definite argument NP of the unaccusative verb to occur postverbally" (291).

Para responder a la cuestión planteada, debemos tener en cuenta que existe una evidente falta de sistematicidad en el tratamiento de las distribuciones ordenativas que exhiben los predicados monovalentes en español, además de la escasez de estudios sobre la inacusatividad en esta lengua, por lo que para abordar el fenómeno del orden VS motivado por los tres factores analizados, consideramos que la opción más fácil y adecuada es recurrir a la memorización, dado que esta es una de las características esenciales de la tradición pedagógica de nuestro ámbito educativo:

> Ciertamente, y como queda demostrado en varios trabajos, el estudiante chino presenta una tendencia bastante generalizada a basar una buena parte de su estudio en la memorización. La lectu-

> ra que se ha hecho de su validez como técnica de aprendizaje ha pasado por varias fases a lo largo de las diferentes épocas. En lo que respecta a nosotros, somos partidarios de la concepción que tienen de ella Biggs (1994, 1996), Dahlin y Watkins (2000) y Sánchez Griñán (2009). Estos autores entienden que los aprendientes chinos emplean estrategias cognitivas profundas al mismo tiempo que repiten, recitan o memorizan. Así pues, la memorización queda al servicio de la interiorización […] El papel predominante del estudiante en el aula es el de escuchar, observar atentamente y hacer el esfuerzo necesario para memorizar todo lo que se dice. No solo memoriza lo que cuenta el docente si no lo que aparece en los materiales con los que se trabajan. Los estudiantes suelen aprenderse *de memoria* dichos materiales. (López Tapia 36)

En nuestro caso específico aplicaremos la memorización proporcionando a nuestros alumnos una tabla, su complejidad se adaptará al nivel de estos, en la que se recogerán los verbos ergativos tendentes al orden VS, para ello se puede utilizar la taxonomía semántica realizada por Hatcher[18] (1956) según la cual clasificó este tipo de verbos como sigue: existencia-presencia, ausencia, comienzo, continuidad, movimiento, verbos de estado puro, verbos de sonido, verbos de color, brillo y luminosidad, verbos de proceso y verbos de desaparición. Con el fin de no extender y complicar la tabla en exceso, de cada conjunto se puede realizar una selección de verbos en función de la frecuencia [19], sea en chino o en español. A continuación se ofrece una muestra de cómo se podría confeccionar esta tabla, todos los ejemplos de la misma son inventados[20]:

ESPAÑOL Orden VS	CHINO Orden VS	INGLÉS Orden SV
VERBOS DE EXISTENCIA-PRESENCIA: *existir*		
Existen muchas constelaciones	有很多星座 Pred.: 有 Suj.: 很多星座 (Pinyin: yǒu hěnduō xīngzuò)	There are many constellations
VERBOS DE AUSENCIA: *faltar*		
Faltan dos estudiantes	沒來兩名學生 Pred.: 沒來 Suj.: 兩名學生 (Pinyin: méi lái liǎng míng xuéshēng)	Two students are missing (for the last one)
VERBOS DE MOVIMIENTO: *llegar*		
Ha llegado un nuevo estudiante	來了一個學生 Pred.: 來了 Suj.: 一個學生 (Pinyin: láile yī xīn xuéshēng)	A new student has arrived
VERBOS DE DESAPARACIÓN: *desaparecer*		
Ha desaparecido un cuadro	消失了一幅畫 Pred.: 消失了 Suj.: 一幅畫 (Pinyin: xiāoshīle yī fú huà)	A painting has disappeared

Tabla 3. Predicados inacusativos con orden VS según Hatcher

Las posibilidades de confeccionar recursos de este tipo son diversas, así como las actividades que se pueden desarrollar a partir de estos, entre otras cuestiones nos van a posibilitar:

1. Tratar las clases verbales sea de forma más general transitivos-intransitivos, inacusativos-inergativos, o de modo más específico y, por ejemplo, enfocarse en los verbos de movimiento y sus diferencias de orden. Los denominados verbos direccionales que "[…] signals the direction in which the subject moves as the result of the displacement" (Li y Thompson 58) tienden a la distribución VS tanto en chino mandarín como en español, ejemplos inventados con dos de estos verbos: 抵達一架

飛機 (pinyin: dǐdá yī jià fēijī, trad. *Llega un avión*) y 來了颱風 (pinyin: láile táifēng, trad. *Viene un tifón*). Sin embargo, "Not all intransitive verbs of motion, however, allow the noun phrase signaling the entity involved in the motion to be in the postverbal position" (Li y Thompson 517), ejemplos inventados: 飛機在飛 (pinyin: fēijī zài fēi, trad. *El avión vuela*) y 登山者攀登 (pinyin: dēngshān zhě pāndēng, trad. *El montañero escala*). Estos casos también nos permitirían enseñar a nuestros alumnos las diferencias entre los sujetos agentivos y los pacientes, puesto que en los verbos de movimiento que no permiten el orden VS (*saltar, escalar, volar, etc.*) "[…] el sujeto es un caso de macropapel semántico actor, el cual puede ejercer un control sobre la actividad significada por el verbo" (Alonso-Cortés 196).

2. Los SSNN sujetos escuetos permiten *jugar* con la diferenciación morfológica entre sustantivos comunes y propios. Asimismo, la definitud se marca en español las más de las veces mediante el uso de artículos definidos, elementos de los que carece el chino mandarín, de ahí que podamos a su vez abordar la problemática que supone para los estudiantes sinohablantes el uso de los artículos.
3. Por su parte, los SSNN sujetos complejos posibilitan tratar en clase los diferentes tipos de modificadores del sustantivo: adjetivos, complementos del nombre preposicionales, relativos, etc., estos, junto con los mencionados artículos, también suponen un gran escollo pues la lengua china carece de relativos.

4. Conclusión

En este artículo se ha analizado un aspecto del orden de palabras como es la disposición VS no marcada informativamente con ciertos predicados inacusativos, más los factores SSNN sujetos escuetos, relacionados directamente con este

tipo de predicados, y los sujetos complejos como contraparte de los anteriores.

El hecho de haber optado por esta colocación se debe, al margen de que abogamos por introducir fenomenología relativa al orden de constituyentes en nuestras clases de ELE, a que la transferencia de estructuras en este marco educativo no se produce de la L_1-chino mandarín a la lengua meta-español, sino que, como hemos expuesto, se da desde la L_2-inglés por la semejanza que los alumnos perciben y precisamente esta lengua se caracteriza por unos requisitos estructurales más rígidos o menos flexibles en cuanto al orden que el español:

> His natural desire [el del estudiante] to transfer structures directly from English to Spanish may overcome his grasp of important differences. [El estudiante] needs to be guided to understand the grammatical values of lexical items in various contexts, and he needs to be shown the similarities and differences of structure that affect word order. To the extent such guidance can be given in terms of general principles and unified patterns, the student is given greater power over the language, with a corresponding increase in learning efficiency. (Stiehm 430)

Una apropiada comprensión del engranaje ordenativo que poseen las lenguas desencadenará ese *poder sobre la lengua* al que alude Stiehm en el alumnado, puesto que este será capaz de entender que los elementos a nivel sintagmático o clausular no se disponen en función de una mera adición secuencial, sino que obedecen a ciertos principios de estructuración jerárquica que gobiernan la adición de constituyentes y que estos principios, aunque presentes en todas las lenguas, en su realización varían de unas a otras.

Bibliografía

Alonso-Cortés, Ángel. "Rasgos activos y ergativos del español". *Verba* 28 (2001): 187-206. Impreso.

Bosque, Ignacio y Javier Gutiérrez-Rexach. *Fundamentos de Gramática Formal*. Madrid: Akal, 2009. Impreso.

Burzio, Luigi. *Italian Syntax: A Government-Binding Approach*. Dordrecht: Reidel, 1986. Print.

Delbecque, Nicole. *Gramática española: enseñanza e investigación. II, Gramática. 4, El orden de los sintagmas. La posición del regente*. Salamanca: Universidad de Salamanca, 1991. Impreso.

Dik, Simon C. *The Theory of Functional Grammar. Part 1: The Structure of the Clause*. Ed. Kees Hengeveld. Berlín: Mouton de Gruyter, 1997. Print.

Dixon, Robert M. W. "Ergativity". *Language* 55/1 (1979): 59-138. Print.

Corder, Stephen P. "The Significance of Learner's Errors". *International Review of Applied Linguistics in Language Teaching* 5/4 (1967): 161-170. Print.

Fernández Fuertes, Raquel. "English/Spanish Syntactic Contrasts: Minimalism and Optimality". Tesis doctoral. Universidad de Valladolid, 2001. Impreso.

Fernández Soriano, Olga. "Sobre el orden de palabras en español". *Dicenda: Cuadernos de Filología Hispánica* 11 (1993): 113-151. Impreso.

Gili Gaya, Samuel. *Curso superior de sintaxis española*. Barcelona: Vox, 1980. Impreso.

Givón, Thomas. *On Understanding Grammar*. Academic Press: Nueva York, 1979. Print.

Greenberg, Joseph H. "Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements". *Universals of Human Language*. Ed. Joseph H. Greenberg. Cambridge: MIT Press, 1963. 73-113. Print.

Hatcher, Anna G. "Theme and Underlying Question". Supplement to *Word*. Vol. 12, monograph 3, 1956. Print.

Hawkins, John A. *Word Order Universals*. Nueva York: Academic Press, 1983. Print.

Herschensohn, Julia. *The Second Time Around: Minimalism and L2 Acquisition*. Amsterdam: John Benjamins, 2000. Print.

Hertel, Tammy J. "Lexical and Discourse Factors in the Second Language Acquisition of Spanish Word Order". *Second Language Research* 19/4 (2003): 273-304. Print.

Levin, Beth y Malka Rappaport-Hovav. *Unaccusativity: at the Lexical Syntax-Semantics Interface*. Cambridge, MA: MIT Press. 1995. Print.

Li, Charles N. y Sara A. Thompson. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press, 1981. Print.

López Meirama, Belén. "Aportaciones de la tipología lingüística a una gramática particular: el concepto orden básico y su aplicación al castellano". *Verba: Anuario galego de filoloxia* 24 (1997a): 45-82. Impreso.

---. *La posición del sujeto en la cláusula monoactancial en español*. Santiago de Compostela: Universidad de Santiago de Compostela, 1997b. Impreso.

López Tapia, Francisco J. "La enseñanza del léxico a estudiantes chinos. Criterios para el análisis del tratamiento del léxico en manuales de E/LE". Trabajo de Fin de Máster. Universidad de Barcelona-IL3, 2011. Impreso.

Lozano, Cristóbal y Amaya Mendikoetxea. "Interface Conditions on Postverbal Subjects: A Corpus Study of L2 English". *Bilingualism: Language and Cognition* 13/4 (2010): 475-497. Impreso.

Melis, Chantal y Milagros Alfonso Vega. "Circunstantes, sujetos y orden de palabras en la oración intransitiva del español". *Lingüística* 29/2 (2013): 127-151. Impreso.

Méndez Marassa, Eduardo. "Problemas de los estudiantes chinos de español. Ejercicios específicos". *Suplementos SinoELE* 1 (2009): 1-99. Impreso.

Mendikoetxea, Amaya. "Construcciones inacusativas y pasivas". *Gramática descriptiva de la lengua española 2*. Eds. Ignacio Bosque y Violeta Demonte. Madrid: Espasa Calpe, 1999a. 1631-1722. Impreso.

---. "Construcciones con *se*: medias, pasivas e impersonales". *Gramática descriptiva de la lengua española 2*. Eds. Ignacio Bosque

y Violeta Demonte. Madrid: Espasa Calpe, 1999b. 1575-1630. Impreso.

Moreno Cabrera, Juan Carlos. *El universo de las lenguas. Clasificación, denominación, situación, tipología, historia y bibliografía de las lenguas*. Madrid: Alianza, 2003. Impreso.

Perlmutter, David M. "Impersonal Passives and the Unaccusative Hypothesis". *Proceedings of the Fourth Annual Meeting of the Berkeley Linguistics Society* (1978): 157-189. Print.

Radford, Andrew. *Minimalist Syntax. Exploring the Structure of English*. Cambridge: Cambridge University Press. 2004. Print.

Schwartz, Bonnie D. y Rex A. Sprouse. "L2 Cognitive States and the Full Transfer/Full Access Model". *Second Language Research* 12/1 (1996): 40-72. Print.

Selinker, Larry. "Interlanguage". *International Review of Applied Linguistics in Language Teaching* 10/1-4 (1972): 209-231. Print.

Stiehm, Bruce G. "Teaching Spanish Word Order". *Hispania* 61/3 (1978): 410-434. Print.

Suárez, Mercedes. "Las propiedades del sujeto y la inacusatividad de los verbos de movimiento en castellano medieval". *Itinerarios* 23 (2016): 155-174. Impreso.

Suñer, Margarita. *Syntax and Semantics of Spanish Presentational Sentence Types*. Washington D.C.: Georgetown UP. 1982. Print.

Watts, Bertha M. "The Teaching of Word Order". *College English* 10 (1949): 328-334. Print.

Yen-hui, Audrey L. *Order and Constituency in Mandarin Chinese*. Dordrecht: Kluwer, 1990. Print.

Yuan, Bopin. "Acquiring the Unaccusative/Unergative Distinction in a Second Language: Evidence from English-speaking Learners of L2 Chinese". *Linguistics* 37/2 (1999): 275-296. Print.

Notas

1 En los casos en los que algún alumno sabe taiwanés, el español queda relegado a una L4.

2 En este trabajo por sinohablantes entendemos hablantes taiwaneses cuya lengua nativa es el chino mandarín.

3 Sobre la existencia de tendencias o preferencias en el orden de constituyentes más que de reglas o preceptos de carácter taxativo véase Gili Gaya, Samuel. *Curso superior de sintaxis española*. Barcelona: Vox, 1980, página 86 y ss.

4 Seguimos la distinción postulada por López Meirama ("Aportaciones" 47 y ss.) en la que el orden básico se concibe como un subtipo del no marcado.

5 Por objeto se entiende el complemento directo, es decir, el patrón básico se limita a verbos bivalentes con los SSNN en función de sujeto y objeto expresos. Para entender por qué es la cláusula principal –no subordinada–, declarativa, de polaridad afirmativa, activa y transitiva la que se asocia exclusivamente al patrón básico véase Givón, Thomas. *On Understanding Grammar*. Academic Press: Nueva York, 1979.

6 Tales beneficios también se producen en la adquisición de la propia L1.

7 En este estudio utilizamos como sinónimos orden de constituyentes y orden de palabras ya que nos centramos únicamente en el nivel clausular.

8 Con respecto a mi labor docente mis alumnos oscilan entre el nivel B1 y B2 de español según el Marco Común Europeo de Referencia para las Lenguas.

9 Este predominio del inglés se ha intensificado desde la instauración del plan bilingüe chino mandarín-inglés por parte del Gobierno taiwanés. Véase más información en el siguiente enlace web https://www.fsc.gov.tw/en/home.jsp?id=253&parentpath=0,4,212,252 [21 jul. 2021].

10 Tipológica, genética y geográficamente las diferencias entre el chino mandarín por un lado y el inglés y español por el otro son ampliamente conocidas. Para ahondar más en este tema véase Moreno Cabrera, Juan Carlos. *El universo de las lenguas. Clasificación, denominación, situación, tipología, historia y bibliografía de las lenguas*. Madrid: Alianza, 2003.

11 De ahí que anteriormente nos hayamos referido al inglés como lengua puente.

12 En función de la posición relativa que pueden adoptar los *meaningful elements* de la cláusula, Greenberg planteó seis posibles órdenes básicos: SVO, SOV, VSO, VOS, OSV y OVS.

13 Pues se debe tener en cuenta que: "[…] Mandarin is a language in which «subject» is not a structurally definable notion. In fact, it seems sensible to regard Mandarin as topic-prominent rather than a sub-

ject-prominent language, since the basic structure of sentences can be more insightfully described in terms of the topic-comment relation rather than in terms of the subject-predicate relation" (Li y Thompson 19).

14 Notación de la teoría X-barra desarrollada por el lingüista estadounidense Jackendoff que indica que el núcleo verbal se ha proyectado: verbo + complemento, es decir, este núcleo ha tomado un complemento, los complementos se diferencian de los adjuntos por una cuestión de rección de sus núcleos sobre estos. Para más información véase Bosque, Ignacio y Javier Gutiérrez-Rexach. *Fundamentos de Gramática Formal*. Madrid: Akal, 2009.

15 La inversión que producen estos circunstanciales, especialmente los de lugar y tiempo, es diferente a la que se produce con ciertos inacusativos o con los SSNN escuetos: "In the case of Spanish, Fernández Ramírez [...] already pointed out that Spanish has a tendency to postpone the subject and have SV inversion. He distinguishes between two types of inversion: absolute inversion which is not the result of the anteposition of other elements; and inversion that is indeed caused by the anteposition of other elements" (Fernández Fuertes 109-110).

16 Para más información véase Radford, Andrew. *Minimalist Syntax. Exploring the Structure of English*. Cambridge: Cambridge University Press, 2004.

17 Agradezco su gentil ayuda a mi profesora de chino mandarín Lo, Shiuan You por revisar las traducciones que he realizado a esta lengua.

18 "Junto al orden A V O, existe en la cláusula intransitiva el orden VINTR. – S. Esta observación aparece por primera vez en el estudio de Anna G. Hatcher (1956), que señaló un conjunto de verbos intransitivos donde el orden normal es VINTR. – S. Hatcher hace una clasificación semántica de los verbos intransitivos que manifiesta este orden" (Alonso-Cortés 192-193).

19 Si bien la frecuencia no es un factor determinante, para la tipología lingüística resulta fundamental: "[…] la tipología lingüística ha hecho de la frecuencia uno de los rasgos más importantes que deben considerarse en la descripción de los fenómenos gramaticales de los que se ha ocupado" (López Meirama, "Aportaciones" 52).

20 A fin de no dificultar la comprensión del alumnado en exceso, los ejemplos que proponemos son oraciones declarativas, afirmativas e independientes.

PITFALLS IN MACHINE TRANSLATION: A STUDY ON WEB-BASED NEWS ARTICLES

TRAMPAS EN LA TRADUCCIÓN AUTOMÁTICA: UN ESTUDIO BASADO EN TEXTOS PERIODÍSTICOS EN LÍNEA

Sara Cecilia Moreno-McMahon
cecymcmahon@msn.com
Yasmine Beale-Rivaya
Professor of Spanish Linguistics, Texas State University
NEH Distinguished Professor in the Humanities
yb10@txstate.edu

Recibido: 7 de septiembre 2021. Aceptado: 1 de diciembre 2021

Abstract: A quality translation involves a process that goes beyond matching words, phrases, and sentences between two languages. Machine Translation (MT) systems use applications to translate text from one language to another without the intervention of a human translator (TH), resulting in poor translations. Through this research, we identified the most frequent errors in translations produced by automatic translation systems from English to Spanish. We analyze MT's limitations regarding grammatical, semantic, and lexical elements in a comparative and statistical analysis. Google Translate™ (GT) and Bing Translator™ (BT) were used to translate twenty news stories from English to Spanish. Both services ensure that, through neural machine translation (TAN), systems can "think" for themselves without human intervention. TAN

systems have achieved impressive results using state-of-the-art sentence processing techniques. Although these systems appear to be very promising, they still produce many errors in their translations.

Keywords/Palabras claves: Automatic translation/Traducción automática, errors in translation/errores de traducción, post-edition/posedición, journalism in Spanish/periodismo en español, Spanish in the United States/español de Los Estados Unidos.

Resumen: Una traducción de calidad conlleva un proceso que va más allá de emparejar palabras, frases y oraciones entre dos idiomas. Los sistemas de traducción automática (TA) utilizan aplicaciones para traducir un texto de un lenguaje a otro sin la intervención de un traductor humano (TH) y, como resultado producen traducciones poco satisfactorias. Mediante esta investigación, identificamos los errores más frecuentes en las traducciones que producen los sistemas de TA del inglés al español. Analizamos las limitaciones de la TA en lo que respecta a los elementos gramaticales, semánticos y léxicos en un análisis comparativo y estadístico. Google Translate (GT) y Bing Translator™ (BT) sirvieron para traducir veinte noticias periodísticas del inglés al español. Ambos servicios aseguran que, a través de la traducción automática neuronal (TAN), los sistemas pueden "pensar" por si mismos sin la intervención humana. Los sistemas de TAN han logrado impresionantes resultados mediante técnicas de procesamiento de oraciones con la última tecnología de vanguardia. Aunque estos sistemas aparentan ser muy prometedores, aun producen un gran número de errores en sus traducciones.

1. Introducción

La traducción automática (TA) se ha convertido en una herramienta indispensable entre las agencias de traducción y las grandes empresas internacionales. Para facilitar el trabajo del traductor humano (TH) y, para aliviar la demanda

que requiere el mercado internacional, se han implementado procesos para la internacionalización de la comunicación digital. Existe una plétora de servicios de traducciones que han mejorado su tecnología para proporcionar los beneficios de productividad que las empresas requieren. Las compañías que operan las traducciones automáticas hacen posible la traducción en poco tiempo y a bajo costo. Según el sitio de traductores Proz.com, el costo promedio de traducciones por traductores humanos (TH) es de diez centavos por palabra o veinticinco dólares por hora para las traducciones del inglés al español. El servicio de traducción de Google es gratuito para los usuarios con un límite de hasta cinco mil palabras. Para las traducciones en grandes cantidades de texto, Google ofrece el servicio de traducciones con mensualidades de 20 dólares, lo cual facilita la traducción de hasta un millón de palabras por mes[1]. El bajo costo de las traducciones automáticas pone en desventaja al TH, dado que la mayoría de los usuarios desconocen si los resultados de las traducciones son de calidad. El propósito de esta investigación es identificar los errores más frecuentes en las traducciones que producen los sistemas de TA del inglés al español. Analizamos las limitaciones de la TA en lo que respecta a los elementos gramaticales, semánticos y léxicos en un análisis comparativo entre la TA y el TH.

Pese a los beneficios de la TA, el proceso de traducción puede ser muy complejo, si se toma en cuenta que muchas palabras tienen múltiples significados, las oraciones pueden tener varias interpretaciones posibles y, ciertas estructuras gramaticales en un idioma pueden no existir o traducirse claramente a otro. Además, existen factores extralingüísticos a considerar en una traducción exitosa, como un amplio conocimiento de la cultura y el contexto. El proceso para una traducción de especialidad requiere ciertos aspectos como el estilo, la terminología y, en muchas ocasiones guías de estilo y una investigación minuciosa. Los textos de especialidad, como el periodismo, exigen ciertos criterios en la redacción. La función principal de los géneros periodísticos es de infor-

mar y la información puede transmitirse en diferentes estilos o con diferente intención. El objetivo siempre tiene que ser el mismo: fácil de entender para los lectores, especialmente en un contexto multilingüe como lo es los Estados Unidos (Guerrero 126).[2]

Las características que distinguen el español estadounidense de otros países monolingües son las variedades morfosintácticas y fonológicas de los distintitos países que forman la población de hispanohablantes en los Estados Unidos (Jiménez 45). Lo que estas variedades del español tienen en común es una indiscutible influencia del inglés, que se manifiesta según el tiempo de residencia del hablante en los Estados Unidos. A nivel morfológico, la prefijación y sufijación nominal se presenta en las combinaciones léxicas derivadas del inglés adaptadas al español, como *lunch box* / *lonchera*, o *truck driver* / *troquero*. Las formas infinitivas del inglés al sistema verbal español resultan en conjugaciones como, *to leak* / *liquear*; *to print* / *printear; to quit* / *cuitear* (Pacchiarotti 165).

Los calcos semánticos dan un nuevo significado a palabras españolas en virtud de su semejanza con el inglés, como, *application* / aplicación; *registration* / registración. También se forman dobletes léxicos en la lengua receptora, dado que adquieren el significado de la lengua donante y se sobreponen con términos que tienen el mismo significado en español como, atender / asistir; aplicación / solicitud (Pacchiarotti,165). Los hispanohablantes en Estados Unidos también usan ciertas frases inglesas, que eliminan la forma equivalente del español al incluir calcos sintácticos como "no son tus negocios", del inglés *it's none of your business* (Torres 412). John Lipsky explica que este tipo de calcos no viola ninguna regla sintáctica o de selección léxica del español, sino que se incorpora en el repertorio de modismos regionales (239).

En los Estados Unidos no existe una lengua oficial dado que, la Constitución garantiza el derecho a sus ciudadanos y residentes a usar cualquier lengua. A pesar del gran crecimiento demográfico y económico, el español aún no se ha

impuesto en el país por razones de prestigio social, científico y cultural como en el caso de la lengua inglesa que goza del apoyo y el prestigio en el medio lingüístico (Marcos Marín, 112). Es por esta razón, que se encuentran errores frecuentes en los textos redactados en español debido a la falta de conciencia de la población y de una norma que garantice una comunicación entre sus usuarios (113).

Las noticias disponibles que se redactan originalmente en español, tales como Telemundo o Univisión, son ejemplos de las pocas fuentes informativas disponibles para los hispanohablantes. A pesar de la creciente influencia de los hispanohablantes en EE. UU., la cobertura de noticias de la nación tiende a servir a personas que se comunican principalmente en inglés.[3] La población hispana de EE. UU. alcanzó un récord de 60,6 millones en 2019, un aumento de 930.000 en comparación al año anterior[4], según las estimaciones de población de la Oficina del Censo de EE. UU. recientemente publicadas (Bustamante sin pág.). Según el reporte anual de 2019 del Instituto Cervantes la población en EE. UU. es de 328,2 millones de personas. El 17,8 % se identifica como hispano y, de este grupo el 71,6 % habla español en sus hogares. Las proyecciones del informe 2019 del Instituto Cervantes indican que, en el 2060, EE. UU. será el segundo país hispanohablante del mundo, después de México, y casi uno de cada tres estadounidenses será hispano. Cerca de 22 millones de alumnos estudian español como segundo idioma según datos del año 2019 (Fernández Vítores, David, y García Montero 10).

Pese a la demanda, los medios de comunicación no cuentan con fuentes de noticias suficientes para los hablantes del idioma español en EE. UU. Los resultados de una investigación realizada por Jorge Ignacio Covarrubias muestran que la mayoría de los medios de noticias tanto impresos como digitales son distribuidos gratuitamente y contienen un gran número de anuncios de propaganda para poder subsistir (7). Covarrubias también encontró que más del 60 % de los periodistas consultados respondieron que los periódicos en es-

pañol contribuyen a la conservación del idioma, y una gran mayoría de los encuestados prefieren las noticias en español. De los 93 participantes que respondieron específicamente a la pregunta: "Para los hispanos, ¿conviene más ofrecerles medios en español o bilingües?", el 61.29 % prefiere las noticias en español, el 30 % bilingües y el 8 % cualquiera de las dos (7-41).

La introducción de la tecnología digital en la comunicación ha fomentado el incremento de neologismos cibernéticos en inglés adaptados al español. El español de Estados Unidos gira en torno a la diglosia que se ajusta a la compleja diversidad del español estadounidense. El español informal se encuentra con mayor frecuencia en las redes sociales. Por otro lado, el español formal sirve para la comunicación en contextos gubernamentales, instituciones y artículos periodísticos.

Según Covarrubias, "Los medios digitales representan una proporción ligeramente mayor al de los medios únicamente impresos". Lo que significa que, con el auge de la información digital y el fácil acceso a los dispositivos electrónicos, los medios de comunicación dependen de la TA en línea para hacer llegar las noticias al mundo (6). Gracias a la ayuda de los servicios gratuitos de TA, las noticias se pueden traducir a una infinidad de idiomas en un instante. Sin embargo, estas traducciones son un ejemplo de los retos que aún existen dentro de la practicidad y velocidad de esta tecnología.

2. Estado actual de la TA:

La TA fue una de las primeras aplicaciones creadas en las computadoras para descifrar códigos durante la Segunda Guerra Mundial. El Memorando Weaver, creado por el matemático Warren Weaver, marcó la pauta para la investigación en la TA (Bowker y Buitrago 37-38). Desde el inicio de las traducciones automáticas en los años cincuenta, los expertos en inteligencia artificial aseguraban que la traducción se podría automatizar para que estas máquinas pudieran traducir tex-

tos técnicos. Por consiguiente, podrían reducir el costo de las traducciones principalmente para las empresas, organismos internacionales y servicios de inteligencia militar. Los resultados no fueron obviamente muy satisfactorios dado que los sistemas no estaban lo suficientemente desarrollados como para reproducir los procesos similares al de un humano. Una traducción de calidad implica la capacidad de reproducir procesos cognitivos muy complejos y difíciles de crear, sin importar los avances y mejoras de la tecnología (Oliver 28).

La TA estuvo dominada por los sistemas basados en conocimientos lingüísticos en sus inicios. En la década de los noventa se implantaron las memorias de traducción y los sistemas estadísticos de traducción que consistían en extraer el conocimiento de corpus paralelos. Los investigadores centraron su atención en el desarrollo de herramientas de traducción asistidas por computadora, como son los sistemas de memoria de traducción. Las herramientas de Traducción Asistida por Computadora (TAC) no tenían como objetivo realizar la tarea de traducción. Su función era ayudar a los traductores profesionales en tareas automatizadas, como la búsqueda de diccionarios o comparar sus traducciones con los textos traducidos previamente. Poco a poco, la traducción automática y los sistemas de memoria se fueron integrando a los procesos de traducción y, por ende, permitieron a los traductores profesionales hacer uso de ambos tipos de tecnología (Bowker y Buitrago 38).

Hoy en día, el modelo neuronal es la tecnología de vanguardia de la traducción automática. Un modelo neuronal es un tipo de modelo estadístico formado por un conjunto de unidades de proceso simple, densamente conectadas entre sí (Casacuberta y Peris 66). Estos sistemas representan de forma numérica las palabras y las frases. Con estos sistemas se utilizan las técnicas del aprendizaje profundo, cuya función consiste en imitar el cerebro humano para realizar funciones cognitivas, como la resolución de problemas y el aprendizaje. Un humano adquiere conocimiento a través de la experiencia y, de esa misma manera las neuronas artificiales lo adquieren (Chen 6).

Algunas compañías se han sumado a la lista de servidores

de la TA como Google, Amazon y Bing, en una contienda por conseguir la tecnología más avanzada. La página de servicios de traducciones Amazon Translate™ explica los beneficios de la nueva tecnología de TA a base de redes neuronales (RN). Según Amazon, las RN son una forma de automatización para ofrecer traducciones más naturales y precisas que los algoritmos estadísticos tradicionales y, de traducción basados en reglas. Además, presenta la TA como una alternativa más eficaz que la TH. Amazon Translate™ asegura que los sistemas de TA son la mejor opción para empresas con miras a la internacionalización de páginas web: "Para los equipos de traductores humanos resulta muy difícil ponerse al día con un contenido dinámico. Con Amazon Translate™, se pueden traducir fácilmente enormes volúmenes de contenido generado por el usuario al instante."

El director de ingeniería de GT, Macduff Hughes, afirmó que, en el 2016 se produjo un gran cambio en el que Google incorporó también la traducción automática a base de RN. A diferencia del antiguo método que funcionaba traduciendo sintagmas y palabra por palabra, ahora traduce la oración completa. La simplicidad de este sistema se encuentra en que no es necesario considerar la sintaxis de la LM. Con el nuevo método, Google puede utilizar los millones de traducciones existentes en la web y así seleccionar la traducción más acertada. La calidad de la traducción depende de la cantidad de traducciones con que la RN esté entrenada y condicionada para hacerlo. Al introducir un nuevo idioma al sistema de RN, es fundamental que existan en la web los datos fiables para entrenar al sistema (Rubio sin pág.).

3. Método del estudio

Los métodos para la evaluación automática consisten en sistemas métricos que pueden comparar los resultados de la TA con la traducción profesional de un humano. Las evaluaciones de paridad entre humano-máquina, por lo general

se miden exclusivamente a nivel de oración, debido a que la mayoría de los sistemas de TA no consideran el contexto más allá de los límites de las oraciones. (Laubli, 656). Al seleccionar los criterios de evaluación y a través de una comparación lingüística pudimos determinar si los errores y ambigüedades alteran el mensaje original, o si las traducciones resultan coherentes. Las aplicaciones de GT y BT en los dispositivos móvil facilitaron el proceso de traducción de los textos, con un promedio de 1500 palabras por cada artículo de noticias, suficientes para una evaluación comprensiva.

Tanto los textos periodísticos originales en inglés como su correspondiente traducción al español en GT y BT se introdujeron a una hoja de cálculo para formar un corpus de análisis paralelo. Pudimos identificar y clasificar en los textos traducidos los errores según el tipo y, de esta manera, se puede examinar si existe una diferencia entre las traducciones de Bing y Google.

4. Evaluación cuantitativa

Los resultados de esta investigación muestran que los errores identificados en GT y BT son muy similares. Estos errores, obtenidos de veinte artículos periodísticos, se clasificaron en siete categorías. De acuerdo con esta clasificación, se analizaron las diferencias entre GT y BT. Se realizó la prueba de CHI-Cuadrada de homogeneidad con la hipótesis de investigación (o alternativa) para determinar si existen diferencias significativas entre los errores de las dos plataformas de TA. Como vemos en la tabla 1 y en la figura 1, la frecuencia de los errores muestra que existen pequeñas diferencias entre GT y BT con respecto a los errores por categoría. En las tablas A y B podemos observar que las diferencias observadas en la frecuencia de errores en ambas plataformas no son significativas. Se puede interpretar que ambas plataformas presentan la misma frecuencia de errores en las 20 historias analizadas.

Tabla 1. Frecuencia de errores identificados en GT y BT por categoría.

	Google Translate	Bing Translator
Ambigüedad	40	35
Calco léxico literal	315	324
Calcos estructurales	117	117
Calco semántico	52	53
Falsos amigos	34	44
Frase idiomática	23	22
Traducción incompleta	18	22
Total	**599**	**617**

Figura 1. Comparación de la frecuencia de errores de Google y Bing.

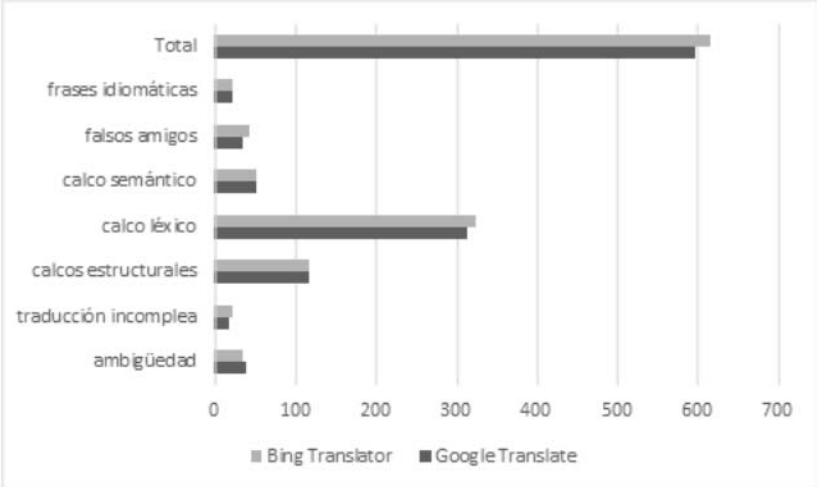


Tabla A. Análisis de homogeneidad con Chi cuadrado (X^2).

Frecuencia observada			
	Google	**Bing**	
Ambigüedad	40	35	75
Traducción incompleta	18	22	40
Calcos estructurales	117	117	234
Calco léxico	315	324	639
Calco semántico	52	53	105
Falsos amigos	34	44	78
Frases idiomáticas	23	22	45
	599	617	1216
Frecuencia esperada			
	Google	**Bing**	
Ambigüedad	36.89154569	38.10845431	75
Traducción incompleta	19.67549103	20.32450897	40
Calcos estructurales	115.1016225	118.8983775	234
Calco léxico	314.3159693	324.6840307	639
Calco semántico	51.64816396	53.35183604	105
Falsos amigos	38.36720751	39.63279249	78
Frases idiomáticas	22.16694079	22.83305921	45
	576	595	1171

Tabla B. Prueba de homogeneidad.

	Google		**Bing**	
	Observada	Esperada	Observada	Esperada
Ambigüedad	40	36.89154569	35	38.10845431
Traducción incompleta	18	19.67549103	22	20.32450897
Calcos estructurales	117	115.1016225	117	118.8983775
Calco léxico	315	314.3159693	324	324.6840307
Calco semántico	52	51.64816396	53	53.35183604
Falsos amigos	34	38.36720751	44	39.63279249
Frases idiomáticas	23	22.16694079	22	22.83305921
Resultado: 1.841368				

En la tabla 2, se observan los resultados en un esquema más detallado con los datos obtenidos de las traducciones de GT. No existen diferencias significativas entre las dos plataformas, y por lo tanto se analizaron solo los datos de Google para presentar estadísticas descriptivas de los errores en 20 artículos de noticias.

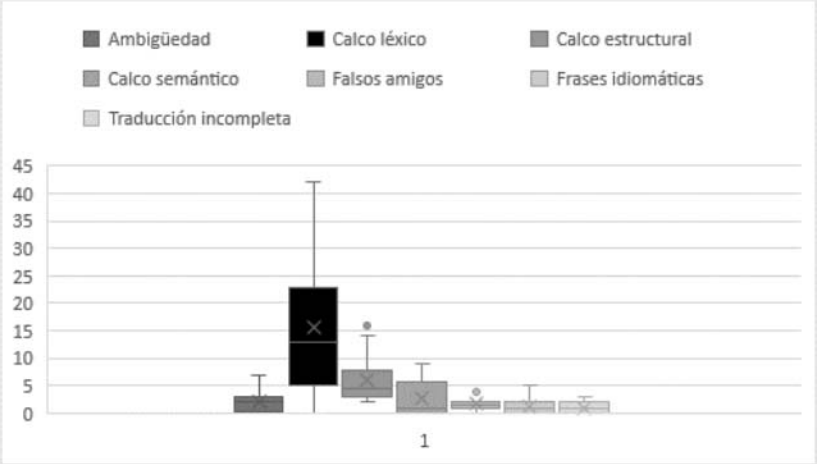
Tabla 2. Frecuencia de errores en GT por categoría y por artículo.

	Ambigüedad	Calco léxico	Calco estructural	Calco semántico	Falsos amigos	Frases idiomáticas	Traducción incompleta	Total
1	1	5	4	2	1	0	2	15
2	1	20	2	1	2	0	1	27
3	7	22	7	1	1	5	2	45
4	2	12	2	0	1	0	1	18
5	2	11	5	9	2	1	1	31
6	1	13	5	7	2	2	3	33
7	3	5	3	0	0	1	0	12
8	0	2	3	0	2	0	0	7
9	2	0	3	5	0	0	2	12
10	2	21	11	6	4	0	0	44
11	6	35	16	1	2	2	2	64
12	4	42	14	6	1	4	0	71
13	3	32	8	5	4	0	1	53
14	0	5	7	0	0	0	1	13
15	2	3	4	1	4	0	0	14
16	0	6	9	0	4	2	0	21
17	0	13	6	0	2	1	0	22
18	3	23	2	0	1	1	0	30
19	1	22	2	6	1	1	0	33
20	0	23	4	2	0	2	3	34
	40	**315**	**117**	**52**	**34**	**22**	**18**	**599**

El análisis de una sola plataforma de traducción nos permite observar los resultados de cada categoría con más de-

talle. La figura 2 muestra la distribución de errores representados en cajas de acuerdo con el número de errores. La categoría de calco léxico cuenta con el mayor número de errores, mientras que la categoría de traducción incompleta está representada con el menor número de errores. Esto también se puede observar en los totales por categorías en la tabla 2. Independientemente de las categorías, todas las observaciones del número de errores son sesgadas y por lo tanto asimétricas. Es decir, la mayoría de las historias tienden a tener un número bajo de errores mientras que uno o dos artículos periodísticos acumularon un mayor número de errores. Por ejemplo, en la historia 12 de la tabla 2 en la categoría del calco léxico se encontraron 42 errores.

Figura 2. Distribución de valores en las categorías de errores.



En la figura 2, la barra horizontal en medio de la caja representa la mediana y la x la media. Se observa en forma cuantitativa y consistente con los diagramas de las distribuciones que la categoría de calco léxico tiene un promedio mayor de número de errores por historia (una mediana de 13 errores por historia). Esta categoría tiene en promedio la dispersión más grande del número de errores (desviación estándar de

11.8 errores). Las demás categorías tienen promedios (medianas) que van de 1 a 5 errores por historia con poca dispersión en comparación con otras categorías. La distribución en una escala horizontal no es simétrica y la mayoría de los errores se encuentran en las categorías que tienden a tener un valor relativamente menor (categorías en azul). La categoría de calcos léxicos (en negro) pertenece a la minoría en el diagrama horizontal, pero con un valor relativamente mayor.

La tabla 3 indica que el promedio de errores por historia en esta categoría es de cerca de 16 errores y, la mayoría de las historias tienen menos de 22 errores. La tabla 3 presenta las distribuciones y los resultados cuantitativos del análisis con las estadísticas de media central y de dispersión del número de errores por categoría. Hemos observado que los valores altos alteran la media de una manera significativa sin dar una medida adecuada del centro de la distribución.

Tabla 3. Comparación de la media, mediana y desviación estándar.

Categoría	**Media**	**Mediana**	**Desviación estándar**
Ambigüedad	2.00	2	1.9
Calco léxico	15.75	13	11.8
Calcos estructurales	5.85	5	4.03
Calco semántico	2.60	1	2.95
Falsos amigos	1.70	2	1.38
Frase idiomática	1.10	1	1.41
Traducción incompleta	0.95	1	1.05

Aunque la categoría de calcos léxicos sobresale en la frecuencia de errores, en una distribución estadística podemos identificar las diferencias de valores entre todas las categorías.

5. Evaluación y análisis

Para interpretar el significado de un texto, un traductor

analiza el mensaje de la lengua fuente (LF) y todos sus componentes, incluyendo un conocimiento amplio de la gramática, sintaxis y semántica de la lengua meta. El número de errores en las traducciones al español y los resultados mostraron que tanto GT como BT presentaron serios errores semánticos y sintácticos que afectaron la calidad y fidelidad de las traducciones. Los resultados muestran que los sistemas de TA actuales tienen muchas limitaciones al traducir oraciones complejas y, por consiguiente, tienden a transferir la estructura y el léxico del inglés al español, con resultados muy literales.

Al evaluar la calidad de los textos traducidos, encontramos que el principal problema se encuentra en los calcos léxicos y los falsos amigos por el grado de semejanza de las palabras entre el inglés y el español en los errores del vocabulario. La categoría de calcos estructurales muestra las formas inglesas que se adaptan al español incluyendo la omisión del artículo, la estructura de la oración, el tiempo verbal, errores en el uso de preposiciones o incoherencias que cambian el significado de las frases.

La tabla 4 representa los datos organizados según la categoría y el número de errores. A su vez, estos se dividieron en dos partes: la primera representa el grupo de los errores relacionados con la semántica, que lo forman las irregularidades en el uso de los cognados falsos, la ambigüedad y lo calcos léxicos. El segundo grupo de la tabla 4 enumera el uso incorrecto u omisión de artículos y preposiciones. Se identificaron algunos errores producidos por la adaptación a la normativa en inglés, tales como el abuso del pretérito pluscuamperfecto, concordancia de sujeto y verbo y, en algunos casos, la concordancia del tiempo verbal y de género.

Tabla 4. Frecuencia de errores y categorías.

Errores semánticos	**Google Translate**	**Bing Translator**
Ambigüedad	40	35

Calco léxico literal	315	324
Calco semántico	52	53
falsos amigos	34	44
Traducción incompleta	18	22
Calcos estructurales		
construcción verbal	1	2
concordancia de género	10	7
frase idiomática	23	24
falta del artículo	14	13
preposiciones	7	7
tiempo verbal	10	12
estructura oracional	52	52

La evaluación de los 20 artículos seleccionados demuestra que tanto Google como Bing produjeron traducciones muy similares: 599 errores en GT y 617 en BT en total. El mayor número de errores pertenece a la categoría de calco léxico literal: GT con 315 errores y BT con 324. Ambas aplicaciones fallaron en la selección apropiada del significado del léxico. Por esta razón, los resultados sugieren que los sistemas de TA no son capaces de traducir algunas de las palabras que presentan polisemia. Debido a que los sistemas de TA operan a nivel oracional, sigue siendo un reto emparejar la semántica del léxico con el contexto.

En la Tabla 5, presentamos un cuadro comparativo de las traducciones de la TA (GT y BT) y la de un traductor humano (TH) y la frecuencia de errores.

Tabla 5. Ejemplo de error según categoría.

Calco léxico	LF	he couldn't **afford** to go to the doctor without insurance
	TA	no podía **permitirse** ir al médico sin seguro
	TH	no podía **darse el lujo** de ir al médico sin seguro
Calco semántico	LF	Canada is a top desired destination for **would-be emigrant** Americans.
	TA	Canadá es un destino ideal para **los aspirantes a estadounidenses emigrantes.**
	TH	Canadá es uno de los destinos principales para **los estadounidenses que desean emigrar a otros países**.
Falsos amigos	LF	Nominated in 1993 by former President Bill Clinton, the liberal **justice** was the legal force behind many of the successes of the women's movement
	TA	Nominada en 1993 por el ex presidente Bill Clinton, **la justicia** liberal fue la fuerza legal detrás de muchos de los éxitos del movimiento de mujeres
	TH	Nominada en 1993 por el expresidente Bill Clinton, **la juez** liberal fue la principal impulsora de los derechos de las mujeres.
Frases idiomáticas	LF	we're not out of the woods yet
	TA	aun no estamos fuera del bosque
	TH	aun no estamos fuera de peligro

Calcos

Se define como calco a la traducción literal de una palabra o sintagma de la LF incorporada a la LM, con o sin adaptación ortográfica. Estas se dividen en dos categorías: el calco léxico y el semántico. Gómez Capuz (2009) propone varias subcategorías del calco léxico, como el literal y el aproximado. Para el propósito de este estudio, emplearemos el calco

léxico literal, que se refiere a la transferencia de una palabra o compuesto nominal de la LF a la LM con un significado literal sin sentido en la lengua receptora e imita la estructura morfológica del modelo de la lengua original (8). Una gran mayoría de los errores en los textos analizados corresponden a las traducciones del calco léxico literal, que son una imitación exacta de palabras o sintagmas de la LF. La tabla contiene algunos ejemplos del calco léxico literal en la traducción.

Tabla 6. Calco léxico literal.

LF	TA	TH
sit-down conversation	conversación sentada	conversación en persona / extendida
he is doing very well	está haciendo muy bien	se siente bien / se siente mejor
walked back that request	caminó hacia atrás esa solicitud	canceló la solicitud
heard back from a staff member	escucharon de un miembro del personal	un empleado respondió
he spent his entire life savings	pasó toda su vida de ahorros	gastó todos sus ahorros

El calco léxico se limita a un contexto oracional. En la tabla 7, podemos observar el calco léxico literal en un contexto más amplio. La frase *school-based*, traducida como *basados en las escuelas* constituye un problema de traducción dado que es una frase inexistente en español. En el ejemplo A3 (tabla 7), el TH puede asociar inmediatamente la frase de la LF con las frases o palabras existentes en la lengua española, en este caso *escolares*. Los calcos ingleses no siempre se traducen de forma adecuada y afectan a la fidelidad que se busca en una traducción.

Tabla 7. Calco léxico literal en contexto.

A1	LF	...as **school-based** nutrition programs serve fewer students, districts are losing money.
A2	TA	...como los programas de nutrición **basados en las escuelas** atienden a menos estudiantes, los distritos están perdiendo dinero.
A3	TH	...como los programas **escolares** de nutrición atienden a menos estudiantes, los distritos están perdiendo dinero.

El siguiente ejemplo muestra los calcos en relación con el sujeto y la construcción verbal y, cómo al infiltrarse la estructura inglesa en la española interfiere con la semántica.

Tabla 8. Calco léxico en construcción verbal.

B1	LF	the president is doing very well.
B2	TA	el presidente estaba haciendo muy bien.
B3	TH	el presidente se siente mejor.

El factor del contexto en el proceso de traducción es crucial. Ambos ejemplos presentan una traducción aislada del contexto. Si incluimos el contexto en el ejemplo B1 de la tabla 8: *The physician tried to tamp down panic, declaring during a news conference Saturday the president was "doing very well,"* la frase *doing very well* puede traducirse de diferentes maneras, según el estilo y contexto. La LF indica que el presidente se siente mejor tras una condición médica. El verbo *hacer* en la LM es irrelevante en este contexto, por lo tanto, el TH en B1 puede explorar las opciones existentes en la LM para su elección: "está mejorando", "se siente mejor", "se está recuperando", entre otras. El ejemplo B2, es un prototipo del calco sintáctico o anglicismo. El ejemplo C2 de la tabla 9 se podría interpretar de diferentes maneras si se considera que, aunque la forma verbal "ha estado" del pretérito perfecto en tercera persona singular no concuerda con el sujeto. Una manera que el TH podría corregir la concordancia verbal sería: "Toda mi vida he estado." O cambiar la estructura completa-

mente sin perjudicar el mensaje original: "He pasado toda mi vida con él." También puede ser "Toda mi vida a girado a su alrededor". La correspondencia que el TH encuentra en *junto a él* concuerda con el significado del texto original.

Tabla 9. Calco léxico literal.

C1	LF	My whole life has been around him.
C2	TA	Toda mi vida ha estado a su alrededor
C3	TH	Toda mi vida la he pasado junto a él. /toda mi vida ha girado en torno a él.

Detectamos las limitaciones de GT en un caso de polisemia de la palabra *left*. Se observa la falta de concordancia verbal entre el léxico y la oración.

Tabla 10. Calco léxico literal y polisemia.

D1	LF	*Devon Austin left with her husband for Costa Rica.*
D2	TA	Devon Austin **izquierda**, con su esposo a Costa Rica.
D3	TH	Devon Austin **se fue** con su esposo a Costa Rica.

El ejemplo D2 (tabla 10) es un caso característico del problema de traducción que con frecuencia se presentó en la categoría de calco léxico literal. La selección errónea de una palabra distorsiona el sentido de la oración y en otros casos, la elección de una palabra parece tener sentido, pero la semántica de la oración cambia completamente según el contexto.

Tabla 11. Calco léxico literal en oraciones.

E1	LF	*It's **overwhelming, she** said. "There are so many people like me **looking".***
E2	TA	**Es abrumador**", dijo. "Hay tanta gente como yo **mirando**".
E3	TH	**Es de preocuparse**, dijo. "Hay tanta gente como yo **buscando trabajo**".

La TA generó una traducción en un concepto aislado don-

de aparentemente el resultado es correcto, (ver E2, tabla 11). E2 parece indicar que algo es tan abrumador y que obliga a alguien a mirar algo. La traducción por un TH (E3) es más específica y, aunque también aislada del contexto no aparenta concordancia con de la oración del inglés. Si a estas oraciones agregamos el contexto original del idioma fuente, el resultado de las traducciones cambia drásticamente.

Tabla 12. Calco léxico literal en contexto.

F1	LF	*Even companies that pay a third of what she once earned won't call her back.* "It's overwhelming," she said. "There are so many people like me **looking**".
F2	TA	Incluso las empresas que pagan un tercio de lo que una vez ganó no la devolverán la llamada. "**Es abrumador**", " dijo. "Hay tanta gente como yo **mirando**."
F3	TH	Incluso las empresas que pagan un tercio de lo que una vez ganó, no le devolverán la llamada. "**Es de preocuparse**", dijo. "Hay tanta gente como yo **buscando trabajo**."

Se observa que un concepto se puede traducir de varias formas, dado que la interpretación de la LF podría dar resultados bastante amplios. Por ejemplo, la palabra *looking* en inglés tiene múltiples significados en español, incluyendo *mirar* y *buscar*. En casos de polisemia, el traductor puede hacer las distinciones semánticas siempre y cuando el resultado de la traducción concuerde con el contexto. El problema de la pérdida léxica en la TA podría ser al mismo tiempo un síntoma o una causa de un problema más grave de los sistemas actuales. Cuando se presenta una oración ambigua en una traducción, el contexto es imprescindible para elegir la traducción que maximice la probabilidad en la claridad en la oración (Vanmassenhove, 1).

Calcos semánticos

A diferencia del calco léxico, el calco semántico modifica el significado de una palabra ya existente en la lengua receptora

y tan solo se limita a alterar los significados de esa palabra. (Gómez Capuz 11). El calco semántico puede aportar un significado parcial o total a la lengua receptora y pueden ser parónimos o no parónimos. Según Gómez Capuz, los calcos semánticos parónimos presentan una similitud formal con palabras de otra lengua que facilita la transferencia semántica, como en el caso de *carpet* y carpeta. Este tipo de calco también se les conoce como "falsos amigos", que veremos en más detalle en la siguiente sección. Los calcos semánticos no parónimos tienen una similitud conceptual entre palabras de dos lenguas, pero no formal ni etimológica, por ejemplo, "perro caliente" y *hot dog*. (14). En la siguiente tabla, se listan ejemplos en donde se detectan los calcos semánticos parónimos.

Tabla 13. Calco semántico (parónimo).

LF	LM	TH
...igniting *protests and violence here*	**. . . e n c e n d i e n d o** protestas y violencia	**...provocando** protestas y violencia
Monica helped ***handle*** *the piles of paperwork.*	Mónica ayudó a **manejar** los montones de papeleo.	Mónica ayudó con el papeleo acumulado.
...were ***struggling*** *with flood damages.*	...estaban **luchando** contra los daños por las inundaciones.	...estaban **sufriendo** por los daños de las inundaciones.
A ***parent*** *with two children*	Un **padre** con dos hijos	Un **padre o una madre** con dos hijos.
swing *states*	los estados **oscilantes**	Estados **indecisos**
Advocates ***target*** *Latinos*	Los defensores **apuntan** a los latinos	Activistas **intentan captar a** los latinos

Falsos amigos

A pesar de las diferencias genealógicas del inglés, una lengua germánica, y el español, una lengua romance, ambos idiomas aún comparten ciertos rasgos de un parentesco (Santos 17). El traductor tiene que poder reconocer las diferencias

semánticas en las similitudes ortográficas entre los cognados falsos y verdaderos. Se consideran cognados verdaderos cuando dos palabras en dos idiomas tienen la misma ortografía, etimología y significado como en los siguientes pares en inglés y español: *patio* y patio *art* y arte, *angel* y ángel. Los falsos amigos (FA) son pares de palabras que tienen una ortografía y pronunciación similar, pero con un significado muy diferente. Estas similitudes pueden conducir a asociaciones falsas, malentendidos, o distorsión del contexto (Vesiburgs, 628). No todos los FA son iguales, dado que se pueden clasificar de diversas maneras.

Estas clasificaciones incluyen a los FA semánticos cuyas palabras son ortográfica o fonológicamente similares en varias lenguas dadas y están relacionadas en su etimología, ya sea porque ambas proceden de una lengua en común como el latín, en *library* y librería. Los FA semánticos a su vez se clasifican en totales y parciales (Vesiburgs, 628). Los FA totales son aquellas palabras cuyos significados son completamente distintos, como en *salad* y salado. Los FA parciales son aquellos que en dos o más lenguas comparten algunos significados mientras difieren en otros (Chamizo 75). Un ejemplo de FA parcial se encuentra en la tabla 5 en la palabra *justice,* y su cognado en español *justicia.* Ambos idiomas comparten los significados de cualidad de ser justo, imparcialidad; pero difieren en el contexto de *justice* como un miembro de la Suprema Corte (*Supreme Court).* Las opciones que concuerdan con el contexto original, como en este ejemplo, la traducción que empareja con la LF es *juez.* La mayoría de los errores identificados en las traducciones dentro de esta categoría, pertenecen al tipo de cognados falsos parciales. Un ejemplo de cognado falso parcial se encontró en la palabra en inglés *move* y su cognado en español mover. Ambas palabras comparten el significado de hacer que un cuerpo deje el lugar y pase a ocupar otro, menear o agitar una cosa o parte del cuerpo. El siguiente ejemplo muestra la disparidad entre la semántica de ambos cognados y cómo el contexto es determinante en la elección de una palabra u otra.

Tabla 14. Falsos amigos.

G1	LF	*Migrant farmworkers, many coronavirus positive,* ***move*** *north from Florida to other states.*
G2	TA	Trabajadores agrícolas migrantes, muchos coronavirus positivos, se **mueven** hacia el norte de Florida a otros estados.
G3	TH	Los trabajadores agrícolas migrantes, muchos de ellos portadores del coronavirus, **se trasladan/mudan** de Florida a otros estados del norte.

Un cognado falso de la palabra en inglés *move* en español puede significar trasladar, mover, mudar, cambiar entre otras (G2 tabla 14). *Mover* no concuerda con la oración de la LF. Aunque *mover* comparte algunas características ortográficas y semánticas con *move*, en el contexto de trasladarse de una región a otra, *mover* pierde su lugar por una palabra que concuerde mejor con el contexto.

Calcos estructurales

Uno de los errores más comunes en el proceso de traducción es la confusión sintáctica en las preposiciones, artículos y la estructura de la oración. Cuando se sigue el modelo de la LF, se crean distorsiones en el significado y la coherencia en la LM. Los calcos identificados en la TA afectan principalmente la estructura sintáctica, preposiciones u omisiones de artículos. En el análisis de los textos se identificaron 117 errores estructurales en GT y BT. Los errores gramaticales que se muestran en la tabla 15 son ejemplos de la transferencia de las estructuras inglesa a la española. La omisión o adición de artículos no concuerda con las reglas de uso españolas. Los calcos estructurales pueden producirse por diferentes mecanismos, como el orden de las palabras que dan como resultado traducciones ambiguas. Aunque estos tipos de errores se consideren menos graves, se han incluido para demostrar las transferencias morfosintácticas de la LF a la LM. Sin embargo, la elección errónea de una preposición puede crear confusión en la semántica, como en el siguiente ejemplo.

Tabla 15. Calco estructural,

H1	LF	*Vasquez needed his father to recover from the coronavirus.*
H2	TA	Vásquez necesitaba **a** su padre para recuperarse del coronavirus
H3	TH	Vásquez necesitaba **que** su padre se recuperara del coronavirus.

El calco del inglés de la proposición *to* al *a* en ocasiones pasa desapercibido. En este caso no se puede ignorar por la ambigüedad que resulta en la traducción. Para resolver este dilema, podríamos preguntarnos si Vásquez es el enfermo que necesita a su padre para poderse recuperar del coronavirus o, si el padre de Vásquez es el enfermo que necesita recuperarse. Aún sin el contexto se puede deducir que la recuperación de Vásquez depende de una intervención médica y no necesariamente de su padre. Veamos el contexto de la LF: "*Vasquez, 40, who has lived with muscular dystrophy since birth, depends on his father to be his arms and legs*". Según el contexto podemos concluir que Vásquez necesitaba que su padre se recuperara por ser el encargado de su cuidado.

Tabla 16. Tipos de calco estructural.

LF	LM	Sugerencia de TH
The eyes are said to be windows to the soul	Se dice que los ojos son ventanas del alma.	Se dice que los ojos son **las** ventanas del alma.
demand limits on voting by mail.	exigen límites **a** votar por correo.	exigen límites **para** votar por correo.
And does it matter if it's for the greater good of the economy?	¿**Y** importa si es por el bien de la economía?	¿E importa si es por el bien de la economía? / y, ¿importa si es por el bien de la economía?
Hernandez's aunt was admitted to the hospital.	La tía de Hernández **fue ingresada** en el hospital.	Internaron a la tía de Hernández en el hospital.

Frases idiomáticas

Tradicionalmente las frases idiomáticas se consideran uno de los factores más idiosincrásicos de una comunidad de hablantes y por lo tanto su traducción puede ser compleja (Molina 429). Según Molina Plaza, en la búsqueda de la correspondencia de las unidades fraseológicas de una lengua a otra no se pueden establecer correspondencias automáticas dado que su complejidad requiere identificar la complejidad de los aspectos léxicos, semánticos, pragmáticos, textuales y discursivos de la expresión idiomática en el contexto de la LM. Todo depende de la dificultad de la frase y de la competencia del traductor (429). Los resultados de este estudio sugieren que las traducciones de GT y BT aún no cuentan con la suficiente competencia para traducir frases idiomáticas del inglés al español. Existen factores culturales y diferencias lingüísticas que influyen en la interpretación y traducción de modismos en las traducciones entre dos lenguas que solo el traductor que conoce ampliamente dichos factores podría descifrar. De los 20 artículos evaluados se identificaron 23 frases idiomáticas, de las cuales BT falló en cada una de ellas, mientras GT solo pudo captar una de ellas.

Table 17. Frases idiomáticas.

LF	TA	TH
business-as-usual	negocios como de costumbre	como si nada hubiera ocurrido
There's no rewinding the clock	No hay forma de rebobinar el reloj	No hay forma de regresar el tiempo
at the drop of a dime	a la caída de un centavo	en un abrir y cerrar de ojos
Justin is cut from the same cloth	Justin esta cortado de una tela similar	Justin está cortado por la misma tijera/ por el mismo patrón

Conclusión

Pese a ciertas ventajas de la TA, es fundamental subrayar que el proceso de posedición beneficia a cualquier traducción realizada automáticamente. Los resultados del análisis sugieren que la TA aún presenta serios errores en la falta de comprensión del contexto más allá de la oración. Las traducciones que producen los motores de TA pueden ser útiles si se incluye el proceso de posedición, ya que los resultados que ofrecen estos sistemas no logran la misma fidelidad que las traducciones revisadas por un TH. La combinación de ambos métodos se ha convertido en el modelo a seguir tanto por traductores profesionales, como por empresas proveedoras de servicios traducción.

La mayoría de los errores se encontraron en las categorías de los calcos léxicos literales, estructurales y semánticos. Se detectó que la elección del léxico fue el principal problema cuando la TA no considera el contexto. Tanto GT como BT fallaron en múltiples ocasiones al no poder solventar los problemas de sinonimia. Las traducciones relacionadas con los textos periodísticos no captaron la esencia y la brevedad que requieren los encabezados. Pudimos identificar que los patrones de selección léxica fueron aleatorios en su mayoría.

Tabla 18. Traducciones de encabezados periodísticos.

LF	Advocates Target Latinos for ACA Enrollment
TA	Los defensores **apuntan** a los latinos para la inscripción de ACA
TH	Activistas se enfocan en los latinos para inscripciones de ACA

El ejemplo de la tabla 18 muestra que el TH toma en cuenta que esta oración es un encabezado para una historia periodística, omite el artículo y selecciona el vocabulario que mejor se adapta al contexto.

Los sistemas de TA carecen de la habilidad de interpretar no solo el contexto, sino formular oraciones con la estructura gramatical española. La función de la traducción automática

neuronal (NMT por sus siglas en inglés) promete que estos sistemas superan a los sistemas predecesores que traducían frase por frase. Los resultados de esta investigación muestran que las traducciones de TA aún son insuficientes para la adecuación y fluidez del texto. Cuando la TA se aplica a las traducciones de textos de especialidad, como los textos periódicos, puede a ser un factor determinante en la calidad de las traducciones.

Representantes de Google afirman que el sistema de redes neuronales (RN) ha sido entrenado con una gran cantidad de traducciones y condicionado para mejorar la calidad. El control de calidad en Google Translate™ se basa en la aportación de usuarios quienes alertan a Google a través de mensajes o en las redes sociales. El sistema no corrige cada traducción incorrecta porque prefieren que el modelo haga su trabajo de "aprendizaje" y en ocasiones tendrán que corregir traducciones que sean ofensivas o engañosas (Rubio sin pág.).

La evaluación de la calidad de la TA ha sido objeto de controversias en la investigación de la industria de servicios de traducciones. Las evaluaciones de traducciones por traductores profesionales son más confiables a pesar de los altos costos que estas requieren. Sin embargo, persiste una tendencia en confiar en estudiantes y evaluadores novatos que a veces muestran un nivel limitado de competencia en los idiomas involucrados. Esto sucede con algunas compañías que ofrecen servicios de TA, ya que los traductores y evaluadores no expertos carecen del conocimiento de traducción y, por lo tanto, pueden no notar sutiles diferencias que hacen que una traducción sea más adecuada que otra. Cuando se enfrentan a una traducción que es difícil de editar, tienden a aceptar la traducción de TA en lugar de intentar mejorarla (Laubli 657).

Existe un mito sobre la traducción automática, al igual que ocurre cuando se plantea la posibilidad de automatizar tareas humanas. Por mucho tiempo, se ha creído que las maquinas reemplazarían a las personas y, por lo tanto, los traductores humanos se convertirían en una especie en ex-

tinción. La TA aún carece de la inteligibilidad, la fidelidad y la precisión requeridas en una traducción de calidad. La comprensión de un texto requiere de un conocimiento muy completo de la lengua de origen, de los mecanismos que procesan la información lingüística y del conocimiento cultural en la lengua fuente y meta (Oliver 31).

El ámbito de la traducción requiere que los traductores profesionales tengan un nivel avanzado no solo en los idiomas, sino un alto conocimiento en la materia de especialidad o la habilidad de investigación. Los últimos avances en la tecnología de la traducción automática prometen que la calidad ha superado sus expectativas y aseguran que la calidad ha llegado a un nivel casi como el de las traducciones realizadas por un humano. La TA, en una era de globalización, se ha convertido en la norma y en una necesidad en la industria de la traducción, pero al mismo tiempo reconocemos que la traducción humana y la evaluación posterior a la TA son indispensables para una eficacia completa.

Works Cited/Obras Citadas

"Average Rates Charged for Translations." *www.Proz.Com*, 2020, https://search.proz.com/employers/rates.

Bustamante, Luis Noe, et al. "U.S. Hispanic Population Surpassed 60 Million in 2019, but Growth Has Slowed." *Pew Research Center*, 7 de julio de 2020.

Bowker, Lynne y Buitrago Ciro, Jairo. *Machine Translation and Global Research: Towards Improved Machine Translation Literacy in the Scholarly Community*. Emerald Group Publishing, 2019.

Chamizo Domínguez, Pedro. "Variaciones representacionales y falsos amigos." *Contrastes*, no. 10, 2005, pp. 73 -103.

Casacuberta, Francisco y Peris, Álvaro. "Traducción automática neuronal". *Revista Tradumàtica*, núm. 15, diciembre de 2017, pp. 66 -72.

Chen, Zhi. "Un estudio empírico del alcance de la traducción automática del español al chino. Caso de estudio de GNMT sobre las expresiones metafóricas y metonímicas". *Círculo de lingüística aplicada a la comunicación*, vol. 83, julio de 2020, pp. 1- 24.

Covarrubias, Jorge Ignacio. "El periodismo en español en los Estados Unidos". Observatorio, marzo de 2016, pp. 1- 86.

Fernández Vítores, David, y García Montero, Luis. El español: una lengua viva, informe 2019. Anuario, Instituto Cervantes, 2019, pp. 1-96.

Guerrero, María José Hernández. "Técnicas específicas de la traducción periodística". *Quaderns*, núm. 13, 2006, pp. 125–39.

Gómez Capuz, Juan. "El tratamiento del préstamo lingüístico y el calco en los textos de bachillerato y en las obras divulgativas". *Tonos: Revista electrónica de estudios filológicos*, núm. 17, julio de 2009, pp. 1-24.

Jiménez, Rosario Outes. "El español escrito en el ámbito educativo de EE.UU.: la acreditación docente en California". *Glosas*, vol. 9, núm. 3, septiembre de 2017, pp. 44–57.

Laubli, Samuel, et al. "A Set of Recommendations for Assessing Human Machine Parityin Lan uage Translation." Journal of Artificial Intelligence Research, vol. 67, mar. 2020, pp. 653–72.

Lipski, John M. "La lengua española en los Estados Unidos: Avanza a la vez que retrocede". *Revista española de lingüística*, vol. 33, núm. 2, 2003, pp. 231–60.

Marcos Marín Francisco A. "RIUSS y el papel de La traducción en el español de los EUA". *II International Colloquiumon Languages, Cultures, Identity, in School and Society*, 2016, pp. 110–18.

Molina Plaza, Silvia. "La traducción de las unidades fraseológicas inglés- español:" *Las palabras del traductor*, 2004, pp. 427–43.

Moreno-Fernández, Francisco. "Español estadounidense: perfiles lingüísticos y sociales". *Glosas*, vol. 9, núm. 2, marzo de 2017, pp. 10–23.

Oliver González, Antoni, et al. *Traducción y tecnologías*. Editorial UOC, 2011.

Rubio, Isabel. "El ingeniero que te permite hablar más de 100 idiomas" | *Tecnología* | *EL PAÍS*, diciembre 2019, https://elpais.com/tecnologia/2019/11/25/actualidad/1574676898_992149.html.

Santos, Pascual, and Aquilino Sánchez. "El inglés y el español desde una perspectiva cuantitativa y distributiva: equivalencias y contrastes." *Estudios ingleses de la Universidad Complutense*, vol. 19, 2011, pp. 15–44.

Vanmassenhove, Eva, et al. "Lost in Translation: Loss and Decay of Linguistic Richness in Machine Translation.", junio de 2019.

Notas

1 Ver "Average Rates Charged for Translations" para más información sobre el costo de traducción.

2 La Academia Norteamericana de la Lengua Española (ANLE) se encarga de determinar las normas del español estadounidense formal, recogiendo como propios aquellos términos que han alcanzado un nivel de uso "culto" ("Estadounidismo no es espanglish"). Algunos de los objetivos principales de la ANLE son: "preservar el uso de la lengua y cuidar el uso de la variante hispano-estadounidense, ofrecer referencia a las instituciones públicas y privadas en el uso de las normas del español de EE. UU. y fortalecer un bilingüismo auténtico que enriquezca la cultura de Estados Unidos" (Academia Norteamericana de la Lengua Española).

3 Ver: Moreno-Fernández, Francisco (2017). Moreno-Fernández propone una realidad del español de Estados Unidos en torno a la diglosia (11). La variedad estadounidense se adquiere de forma oral y en los ambientes familiares donde la presencia del inglés es más evidente con la inclusión de préstamos, calcos, alternancia de códigos y anglicismos tales como la troca o camión; el torque o pavo. La subvariedad culta o general se adquiere a través de la familia, en el ámbito de la educación, contextos comunitarios más generales, en manifestaciones escritas y con menor influencia del inglés. Esta se utiliza en los medios de comunicación social, que coincide con las manifestaciones cultas del español de otros países hispanos (12-13).

4 Ver: Bustamante (2020) con respecto al crecimiento de la población hispana desde 1970 al 2019, según estimaciones de población de la Oficina del Censo de EE. UU.

SECUENCIA GRAMATICAL PARA LA ENSEÑANZA DEL ESPAÑOL COMO LENGUA EXTRANJERA: NUEVAS IDEAS

GRAMMATICAL SEQUENCING IN SPANISH AS A SECOND LANGUAGE: NEW IDEAS

Yun Sil Jeon*, Alejandro Muñoz-Garcés*,
Lindsey Bell and William Jones

Recibido: 15 de septiembre 2021. Aceptado: 1 de diciembre 2021

Resumen: Este artículo es el resultado parcial de una investigación interdisciplinaria en curso sobre Lingüística computacional y la Enseñanza del español como lengua extranjera. La investigación intenta superar la falta de análisis serios sobre estructuras gramaticales en los libros de enseñanza de segundas lenguas, especialmente el orden de las estructuras gramaticales. Empezamos con varios corpus orales en español como C-Or-DiAL, CORLEC y PRESEEA. Más tarde incorporamos al análisis un corpus escrito como el Corpus del Español. Enviamos los corpus a través de un programa computacional de análisis elaborado por nosotros mismos para este fin, que etiqueta las palabras con sus partes del discurso para permitir a los usuarios seleccionar subconjuntos de palabras del corpus, y así determinar la frecuencia en esos corpus. Este análisis nos dará un mejor entendimiento del orden gramatical, porque la secuencia gramatical que plantearemos estará basada en la frecuencia de estas estructuras en los discursos de hablantes nativos, teniendo en cuenta las situaciones comunicativas en los que estas producciones orales y escritas aparecen. Será un paso importante hacia una mejor comprensión de la enseñanza de la gramática, y podría cambiar la forma de enseñarla.

Palabras claves: secuencia gramatical, enseñanza del español, adquisición de lenguas extranjeras.

Abstract: This article is a partial result of an ongoing interdisciplinary research in Computational Linguistics and the Teaching of Spanish as a Second Language. The research intents to overcome the lack of serious analysis on grammar structure of textbooks used for the Second Language Teaching market, especially the order of grammar structures. We started with various oral corpora in Spanish such as C-Or-DiAL, CORLEC, and PRESEEA. In a later stage we incorporated the written corpus, Corpus del Español. We send the corpora through a computational analysis program self-elaborated for this objective in mind, that tag words with their part-of-speech to allows the user to select subsets of corpus words, and determine their frequency in those corpora. This analysis provides a better understanding of the grammar sequencing, as the proposed grammar sequence will be based on the frequency of these structures found in native speaker utterances, taking in account the different communicational situations in which these oral and written productions take place. It will be big step towards a better understanding of the teaching of grammar, and we hope it would change the way of teaching it.

Key words: Grammar sequencing, Teaching Spanish, Second Language Acquisition.

INTRODUCTION: HISTORY OF GRAMMAR SEQUENCE

In the early 1970s, the grammatical skepticism and the growing focus on communicative methods of teaching stole the relevance of grammatical sequencing. Some authorities questioned the belief that it is possible to stablish criteria according to which a grammatical sequence might be organized in a logically, psychologically, and pedagogically satisfactory way.

Since the 1970s, a debate about the primacy of grammatical or functional categories has contributed to the al-

ready prevalent questioning of conventional arrangement of grammatical features. During this period, an attempt was made to give up grammatical sequencing and to build language courses on a semantic basis with grammar taking second place. The belief in the value of some kind of grammatical sequence structuring the teaching of a language has not been lost. However, modern thought on the grammatical syllabus is characterized by selectivity. Grammatical items are not included because they have always been taught, but because they can be justified in relation to the type of course, its length, objectives, and the student population.

Recent experience with notions, functions, and speech acts as well as with discourse and other communicative categories has not provided a replacement for the grammatical syllabus. However, thinking about grammar now tends to be related to functional-communicative aspects of language teaching. The key question is how best to integrate a grammatical syllabus with the demands of a functional or communicative/experiential syllabus. In order to create a grammatical sequence, there is a wide consensus about certain pedagogical principles that have to be kept in mind. Most writers on grammar teaching observe that grammar is cumulative. The concept of requisite knowledge: there are many complex items, which presuppose previous knowledge, and it has to be recognized when creating a grammatical sequence. Finally, no item that has once been introduced should ever be lost sight of.

As a conclusion, it can be said that in spite of doubts about the feasibility of a sequential arrangement, the grammar of a language cannot be taught all at once. Some sort of selection and sequencing is needed when creating a grammatical syllabus for the teaching of a second language. This does not mean that it should be the only organizing principle. It must be related to other subsystems of language as well to the non-linguistic elements –culture for i.e.- that should be part of the second language curriculum.

RESEARCH ORIGIN AND DEVELOPMENT

Grammatical organization and sequence are essential to the success of a language class. The immense majority of professors devoted to the Teaching of Spanish as Second Language follows a method (book or online platform) already in the market, or combines several methods in a class. Following my teaching experience, I used to alter the order of grammatical sequencing, because classes worked better, or because I thought it was more practical. After a while, I started questioning the reason behind the sequence adopted for methods in the market, and why we keep almost the same sequencing even with changes in methods (grammar-translation, direct, audiolingual, the Silent Way, Suggestopedia, community language learning, Total Physical Response, communicative approach, etc). At the same time, I started asking instructors thought of the sequence, and the majority agreed that they were not happy with it, but they used them because it was easy. When I approached editorial people questioning why they do not offer any novelty in sequencing, they responded that they did not see the reason for any changes if instructors were not backing it up: conclusion, it is a vicious circle.

As an attempt to respond to the question above mentioned, I am leading an ongoing collaborative research in Coastal Carolina University (with Dr. Alejandro Muñoz-Garcés from the Department of Languages and Intercultural Studies, Dr. William Jones from the Department of Computer Science, and Dr. Lindsey Bell from the Department of Mathematics and Statistics), since 2015. At the very beginning, we started with the collaboration of Dr. Carlota Nicolás-Martínez from the Università di Firenza. She visited us in 2015 with funds provided by Dean Roberts and Dean Ennis. Both Deans also funded the first Student Assistant for two semesters (Spring 2015 and Fall 2016), to work with Dr. Nicolás-Martínez and I with the data analysis. However, at the end of this collaboration, Dr. Nicolás Martínez and I realized we differed in

methods and objectives, therefore we parted ways in 2017. Dean Roberts has funded two Student Assistants for two semesters (Spring 2017 and Fall 2018), but it was the third Assistant, David Eastman, under Dr. Jones´ supervision, who completed the computational data collection and programming successfully, giving us the tool we use now to extract data presented in this article. Dean Ennis gave us fund for the fourth Student Assistant, who worked during Summer 2018 on the statistical analysis of the data. Unfortunately, we could not come up with a satisfactory outcome with this last assistant, but I am glad we could recruit Dr. Lindsey Bell who agreed on working on the statistical analysis.

The Office of Research Services under the leadership of Dr. Roberts saw the value of our research idea, and proposed us to patent it. The Provisional application was submitted in April 2018, and re-filed in April 2019, and April 2020. We were working to present the Final application before Summer 2021, but because of the COVID-19 situation we have decided to re-file the Provisional application one more year. The Final application will be presented in 2022.

DESCRIPTION OF RESEARCH METHODOLOGY

This study utilizes various oral speech corpora in Spanish: C-Or-DiAL (Corpus Oral Didáctico Anotado Lingüísticamente-Didactic Oral Corpus Linguistically Annotated- with 120 thousand words), CORLEC (Corpus Oral de Referencia del Español Contemporáneo-Reference Spoken Corpus of Contemporary Spanish- 1 million 100 thousand words), and PRESEEA (Proyecto Para el Estudio Sociolingüístico del Español de España y de América-Project for the Sociolinguistic Study of Spanish from Spain and America- Goal of 10 million words). Corpus del Español (Spanish Corpus with 40 thousand words) is the only written speech corpus. The computer program used for analyzing the above corpora is a web-based application written in HTML/Javascript that uses the FeeLing

text analyzer to tag words with their part-of-speech and then allows the user to select subsets of corpus words based on those parts-of-speech and determine their frequency in those corpora. There are options that allow for filtering the results based on certain communicational situations in which these oral and written productions take place (lectures, conferences, conversations in informal settings and work settings, dialogues, interviews). In order to allow for verification of the results, the texts containing the desired words are also available for display.

Current Spanish teaching methods present verb systems as a "skeleton" to organize grammar sequencing around, because verbs present the most difficulty for a learner's acquisition of the language, (verb endings in Spanish are all different depending on the mood, tense, person and number). They also present the grammar in bulk, instead of sequencing it gradually for a learner's better acquisition and assimilation. They introduce all tenses at once (first, second, third person singular and plural), and verb tenses are ordered in a certain way that needs deep analysis and revisions. The computational tool analyzes the native language grammar sequencing of Spanish: it takes native speakers´ oral language data and finds out which grammatical structure is used, in which order, how, and when these structures are utilized. Then, we intent to compare our findings with the grammar structure of methods used for the Second Language Teaching market. As a result, these findings provide a language learning system that includes a much better sequenced program to teach a language. Students will learn the structure of the language based on the frequency of these structures found in native speaker utterances, instead of learning a totally arbitrary grammar sequencing as disclosed via prior language teaching methods. This research includes the presentation of a new grammar sequence to the students to make learning process more accurate and easier.

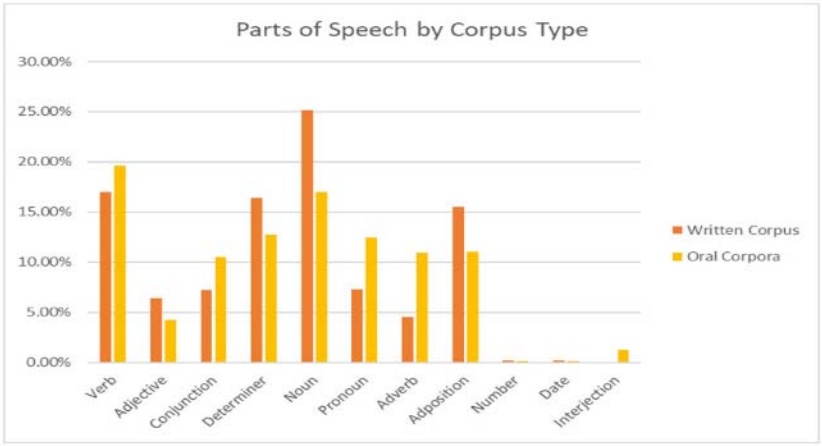
PARCIAL RESULTS

Parts of Speech by Corpus

Verbs are used most frequently among three of the four corpora. Nouns make up the second most common part of speech. A Chi-square test shows that part of speech usage is not homogeneous among the corpora (X-squared = 204548, df = 30, p-value < 0.00001). Post hoc tests reveal in fact that the distribution of parts of speech is different among all four corpora (p-values < 0.00001). This indicates that the four corpora exhibit a variety of utterances of the Spanish language.

Corpus	C-Or-DiAL	Percent	CORLEC	Percent	PRESEEA	Percent	Corpus del Espanol	Percent
Verb	22398	19.84%	152033	19.50%	308543	19.69%	316638	16.99%
Adjective	5592	4.95%	34905	4.48%	63141	4.03%	118798	6.37%
Conjunction	12838	11.37%	73216	9.39%	172721	11.02%	134236	7.20%
Determiner	14321	12.69%	105869	13.58%	192566	12.29%	306042	16.42%
Noun	17889	15.85%	145720	18.69%	254602	16.25%	468311	25.13%
Pronoun	13954	12.36%	85986	11.03%	205485	13.11%	136470	7.32%
Adverb	12857	11.39%	71294	9.14%	185223	11.82%	84856	4.55%
Adposition	12290	10.89%	95346	12.23%	163933	10.46%	289324	15.52%
Number	86	0.08%	1113	0.14%	1401	0.09%	4189	0.22%
Date	93	0.08%	1053	0.14%	1818	0.12%	3643	0.20%
Interjection	579	0.51%	13257	1.70%	17596	1.12%	1263	0.07%
Totals	112897		779792		1567029		1863770	

Corpus	CORLEC	C-Or-DiAL	PRESEEA	Corpus del Espanol
Verb	19.50%	19.84%	19.69%	16.99%
Adjective	4.48%	4.95%	4.03%	6.37%
Conjunction	9.39%	11.37%	11.02%	7.20%
Determiner	13.58%	12.69%	12.29%	16.42%
Noun	18.69%	15.85%	16.25%	25.13%
Pronoun	11.03%	12.36%	13.11%	7.32%
Adverb	9.14%	11.39%	11.82%	4.55%
Adposition	12.23%	10.89%	10.46%	15.52%
Number	0.14%	0.08%	0.09%	0.22%
Date	0.14%	0.08%	0.12%	0.20%
Interjection	1.70%	0.51%	1.12%	0.07%
Totals				



Parts of Speech by Oral vs Written

C-Or-DiAL, CORLEC, and PRESEEA are all oral corpora while Corpus del Español is a written corpus. Therefore, we may consider the usage of parts of speech by corpus type. Chi-squared tests verify that usage of parts of speech is different in written and oral language (X-squared = 189005, df = 10, p-value<0.0001).

Corpus	Oral Copora	Percent	Written Corpus	Percent
Verb	482974	19.64%	316638	16.99%
Adjective	103638	4.21%	118798	6.37%
Conjunction	258775	10.52%	134236	7.20%
Determiner	312756	12.72%	306042	16.42%
Noun	418211	17.00%	468311	25.13%
Pronoun	305425	12.42%	136470	7.32%
Adverb	269374	10.95%	84856	4.55%
Adposition	271569	11.04%	289324	15.52%
Number	2600	0.11%	4189	0.22%
Date	2964	0.12%	3643	0.20%
Interjection	31432	1.28%	1263	0.07%
Totals	2459718	**1**	1863770	**1**

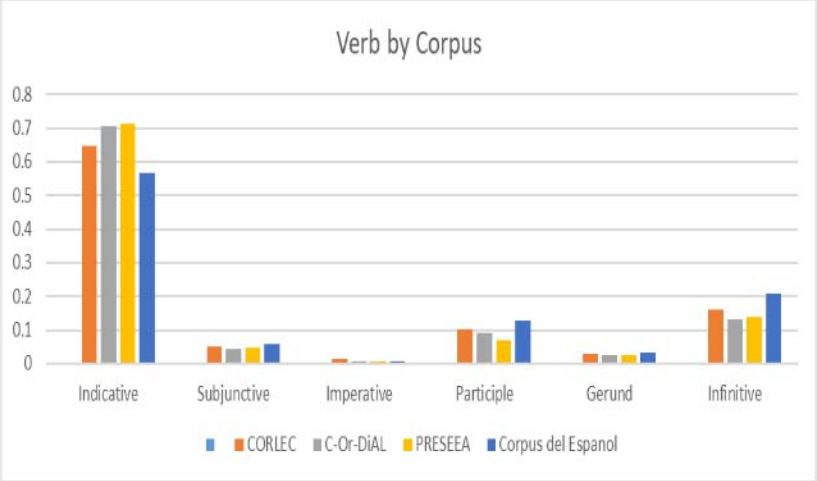
Breakdown of Verbs by Mood and Tense

Among all corpora, indicative present tense is by far the most frequent verb usage at about 46.4%. This is followed by the infinitive at about 17%. Imperative is used the least among the corpora. A Chi-square test shows that verb tense is not homogeneous among the corpora (X-squared =13058 , df = , p-value < 0.00001). Post hoc tests reveal in fact that the distribution of verb tense is different among all four corpora (p-values < 0.00001). The same results are summarized graphically in the following figure.

Verb	C-Or-DiAL	Percent	CORLEC	Percent	PRESEEA	Percent	Corpus del Español	Percent	Overall	Percent
Indicative	15777	70.44%	98446	64.75%	219713	71.21%	179524	56.70%	513460	64.21%
Present	12123	54.13%	81012	53.29%	149089	48.32%	128785	40.67%	371009	46.40%
Imperfect	2268	10.13%	8061	5.30%	34351	11.13%	12538	3.96%	57218	7.16%
Future	163	0.73%	2345	1.54%	1995	0.65%	7341	2.32%	11844	1.48%
Past	1067	4.76%	5262	3.46%	30315	9.83%	26984	8.52%	63628	7.96%
Conditional	156	0.70%	1766	1.16%	3963	1.28%	3876	1.22%	9761	1.22%
Subjunctive	964	4.30%	7938	5.22%	14806	4.80%	18939	5.98%	42647	5.33%
Present	770	3.44%	6605	4.34%	10888	3.53%	15655	4.94%	33918	4.24%
Imperfect	191	0.85%	1328	0.87%	3911	1.27%	3164	1.00%	8594	1.07%
Future	3	0.01%	5	0.00%	7	0.00%	120	0.04%	135	0.02%
Imperative	135	0.60%	1902	1.25%	2293	0.74%	1686	0.53%	6016	0.75%
Participle	2003	8.94%	15349	10.10%	20827	6.75%	40052	12.65%	78231	9.78%
Gerund	542	2.42%	4121	2.71%	8255	2.68%	10680	3.37%	23598	2.95%
Infinitive	2977	13.29%	24277	15.97%	42649	13.82%	65757	20.77%	135660	16.97%
Total	22398	100.00%	152033	100.00%	308543	100.00%	316638	100.00%	799815	100.00%

Verb	CORLEC	C-Or-DiAL	PRESEEA	Corpus del Espanol
Indicative	64.75%	70.44%	71.21%	56.70%
Subjunctive	5.22%	4.30%	4.80%	5.98%
Imperative	1.25%	0.60%	0.74%	0.53%
Participle	10.10%	8.94%	6.75%	12.65%
Gerund	2.71%	2.42%	2.68%	3.37%
Infinitive	15.97%	13.29%	13.82%	20.77%

Firefox about:blank
67 de 149 27/06/2022, 17:25

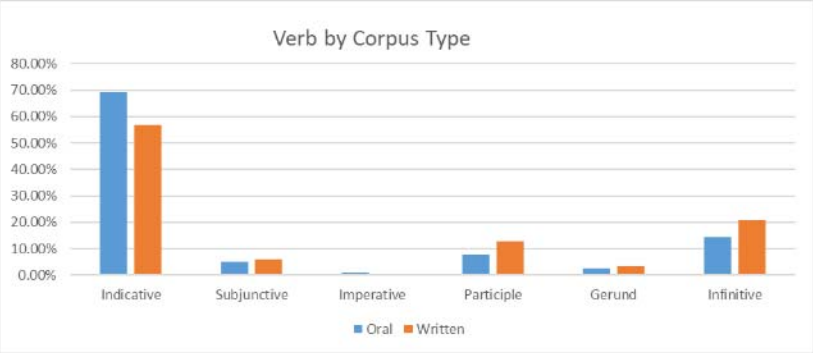


Verbs among Written and Oral by Mood and Tense

We may also consider the distribution of verbs by written and oral language differently. The results are displayed in the following table and figure. There is a difference in verb usage between oral and written language (X-squared = 21171, df = 11, p-value<0.0001).

Verb	Oral Corpora	Percent	Written Corpus	Percent
Indicative	333936	69.14%	179524	56.70%
Present	242224	50.15%	128785	40.67%
Imperfect	44680	9.25%	12538	3.96%
Future	4503	0.93%	7341	2.32%
Past	36644	7.59%	26984	8.52%
Conditional	5885	1.22%	3876	1.22%
Subjunctive	23708	4.91%	18939	5.98%
Present	18263	3.78%	15655	4.94%
Imperfect	5430	1.12%	3164	1.00%
Future	15	0.00%	120	0.04%
Imperative	4330	0.90%	1686	0.53%
Participle	38179	7.90%	40052	12.65%
Gerund	12918	2.67%	10680	3.37%
Infinitive	69903	14.47%	65757	20.77%
Total	482974	100.00%	316638	100.00%

Verb	Oral	Written
Indicative	69.14%	56.70%
Subjunctive	4.91%	5.98%
Imperative	0.90%	0.53%
Participle	7.90%	12.65%
Gerund	2.67%	3.37%
Infinitive	14.47%	20.77%

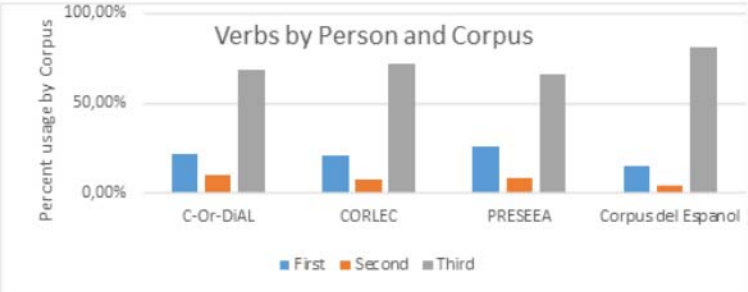


Breakdown of Verbs by Person

A Chi-square test (x-squared = 12658, df =6, p-value <0.0001) show that use of person is dependent on corpus. Post hoc tests show that in fact each corpus is different from the others (p-values <0.0001).

Person	C-Or-DiAL	Percent	CORLEC	Percent	PRESEEA	Percent	Corpus del Espanol	Percent	Overall	Percent
1	3591	21.28%	22780	21.04%	61152	25.82%	30652	15.32%	118175	21.02%
2	1731	10.26%	8119	7.5%	19610	8.28%	7894	3.94%	37354	6.65%
3	11554	68.46%	77387	71.46%	156050	65.90%	161603	80.74%	406594	72.33%
Total	16876	100.00%	108286	100.00%	236812	100.00%	200149	100.00%	562123	100.00%

Firefox about:blank
69 de 149 27/06/2022, 17:25



Nouns in the Oral Corpora

After verbs, nouns are the most important grammatical structure in oral language. To explore themes of the subject of conversation the twenty most frequent nouns were identified in each of the three oral corpora. These are displayed in the following table.

Top 20 Nouns by Frequency in Oral Corpra

Rank	Noun	C-Or-DiAL	Percent	Noun	CORLEC	Percent	Noun	PRESEEA	Percent
1	cosas	195	1.09	día	935	0.64	años	3508	1.38
2	años	195	1.09	verdad	879	0.6	casa	3475	1.36
3	gente	194	1.08	cosas	876	0.6	gente	3368	1.32
4	cosa	184	1.03	años	801	0.55	cosas	2967	1.17
5	verdad	154	0.86	momento	786	0.54	día	2344	0.92
6	ha*	151	0.84	parte	707	0.49	tiempo	1968	0.77
7	trabajo	150	0.84	gente	704	0.48	veces	1940	0.76
8	parte	103	0.58	cosa	702	0.48	vida	1762	0.69
9	periódico	101	0.56	señor	695	0.48	trabajo	1723	0.68
10	momento	96	0.54	tiempo	667	0.46	año	1669	0.66
11	año	94	0.53	casa	622	0.43	uhum	1529	0.6
12	día	92	0.51	año	555	0.38	mamá	1527	0.6
13	vida	90	0.5	vez	536	0.37	cosa	1476	0.58
14	veces	89	0.5	proble-ma	531	0.36	verdad	1435	0.56
15	su*	86	0.48	sí	518	0.36	vez	1402	0.55
16	casa	84	0.47	caso	503	0.35	familia	1387	0.54
17	mundo	83	0.46	mundo	498	0.34	por*	1370	0.54
18	hombre	81	0.45	veces	492	0.34	persona	1334	0.52
19	tiempo	78	0.44	días	462	0.32	barrio	1234	0.48
20	en*	77	0.43	Madrid	458	0.31	amigos	1199	0.47
Totals		17889	13.28		145270	8.88		254602	15.15

Firefox about:blank
70 de 149 27/06/2022, 17:25

Ten nouns were common to the top 20 list among all three corpora. The remaining 10 were missing from at least one corpus. These are highlighted in the following tables.

Nouns Shared by all three Oral Corpra in Top 20				
cosas	casa	día	verdad	gente
años	tiempo	veces	año	cosa

Top 20 Nouns not Shared by All			
Noun	C-Or-DiAL	CORLEC	PRESEEA
ha*	X		
trabajo	X		X
parte	X		X
periódico	X		
momento	X	X	
día	X		X
su*	X		
mundo	X	X	
hombre	X		
en*	X		
señor		X	
vez		X	X
problema		X	
sí		X	
Madrid		X	
uhum			X
mamá			X
familia			X
por*			X
persona			X
barrio			X
amigos			X

PARTIAL CONCLUSIONS

a. Verbs are used most frequently among three of the four corpora. Nouns make up the second most common part of speech. This verifies the idea that instruction based on frequency of word usage in everyday language should be arranged according to verb conjugations with regular introduction of vocabulary or nouns.
b. Grammar sequencing in Spanish teaching textbooks is based on verb conjugation: Present Indicative first, then Preterite Indicative, Imperfect Indicative, Present Subjunctive, and so on. Based on the overall frequency of use among the four corpora, the suggested order of verb introduction follows: Present Indicative, Preterite Indicative, Present Subjunctive and so on. It is clear that early introduction of Present Indicative is important due to the high frequency of use in both written and oral language. We may also consider the distribution of verbs by written and oral language differently. For example, the indicative mood is used about 12% more often in oral language than written language, while infinitive is used about 6% more often in written language than oral. Among all corpora, indicative present tense is by far the most frequent verb usage at about 46.4%. This is followed by the infinitive at about 17%. Imperative is used the least among the corpora.
c. Use of person is dependent on corpus. Post hoc tests show that in fact each corpus is different from the others. However, we can generally see agreement that third person is most commonly used, followed by first person and then second person.
d. Usage of parts of speech is different in written and oral language. For example, interjections, conjunctions, and pronouns are more common in oral language than written language. Verbs are the most

commonly used part of speech in oral language, while nouns are the most commonly used part of speech in written language.

e. After verbs, nouns are the most important grammatical structure in oral language. The summaries suggest the majority of conversation revolves around relationships with other people, daily activities, and time. Percentages given tell us the percent make-up of all the nouns in the corpus.

PLANS FOR FUTURE WORK

a. In oral corpora, verbs are the most commonly used word followed by nouns; while in written corpus, exactly the opposite happens. It is clear that usage of parts of speech differs in written and oral language. We will have to incorporate more written corpora in our analysis to confirm these differences or refute it. Two different grammar sequencing: one for teaching oral skills, and another one for written?
b. If we base the grammar sequencing on verb conjugation, we will need to run more analysis to come up with the best order: what is clear is that it will be different than the grammar sequencing offered by methods in the market. Present tense Indicative Mood is clearly the most frequent, followed by infinitive. However, more analysis is needed to determine if "these infinitives (or participles or gerund)" are part of verbal periphrases or not. Third person is most commonly used, followed by first person and then, second person. This is also something we have to reorder in our new sequence order: not only tenses and mood of verbs, but also person sequence. We will need to refine our computer tool to reflect changes needed.
c. We are researching for finding a way to incorporate

nouns in the analysis, for this to start reflecting different communicational situation that will affect the grammar sequencing order.

This is an ongoing research, which immediate utility is for the teaching Spanish. However, when finished, this could extend to the teaching of other languages. We are confident that this research will become a significant advance, since through these findings we are going to propose a better sequenced grammar. It will be big step towards a better understanding of grammar teaching, because the proposed sequence will be based on the frequency of grammar structures found in native speaker utterances.

BIBLIOGRAPHY

American Council on the Teaching of Foreign Languages (2012). ACTFL Proficiency Guidelines 2012. Virginia. Web.

C-Or-DiAL (2012). Corpus Oral Didáctico Anotado Lingüísticamente. Firenze: Università degli Studi di Firenze. http://www.lllf.uam.es/ESP/Corlec.html Consulted: 2015-Present.

CORLEC (1992). Corpus Oral de Referencia del Español Contemporáneo. Madrid: Laboratorio de Lingüística Informática de la Universidad Autónoma de Madrid. https://e-l.unifi.it/pluginfile.php/434597/mod_book/chapter/1238/C-Or-DiAL_LICEUS.pdf Consulted: 2015-Present.

Corpus del Español (2001). National Endowment for the Humanities. https://www.corpusdelespanol.org/web-dial/ Consulted: 2015-Present.

Goldberg. D., Looney, D., and Lusin, N. (2015) "Enrollment in Languages Other than English in United States Institutions of Higher Education." *Modern Language Association*. Retrieved from https://www.mla.org/content/download/31180/1452509/emB_enrllmnts_nonengl_2013.pdf

Halliday, M.A.K., McIntosh, A. and Strevens, P. (1964). *The Linguistic Sciences and Language Teaching* (pp. 80, 35, 136). London: Longman.

Higgs, Theodore V. (1985) Teaching Grammar for Proficiency. *Foreign Language Annals* 18, 289-96. Print

Jeon, Yun Sil and Alejandro Munoz-Garces (2018). The Grammatical Sequence in the teaching of Spanish as a Second Language in the USA. *Revista Iberoamericana de Lingüística N 13*, Madrid, Spain.

Mackey, W.F. (1965) *Language teaching Analysis* (pp 135). London: Longman.

Miller, William I. (1984) The Scope of the Spanish Grammar for Minimal Communicative Competence. *Foreign Language Annals* 22, 47-53. Print

Moneypenny, Dianne Rosalie Aldrich (2018). "Developing Oral Proficiency in Spanish Across Class Modalities." *CALICO Journal* Vol. 35 Issue: number 3. 257-273. Print.

Moneypenny, D.B., and Aldrich, R.S. (2016) "Online and Face-to-Face Language Learning: A Comparative Analysis of Oral Proficiency in Introductory Spanish. *Journal of Educators Online*, 13(2). 105-133. https://doi.org/10.9743/JEO.2016.2.2

PRESEEA (2014): Corpus del Proyecto para el estudio sociolingüístico del español de España y de América. Alcalá de Henares: Universidad de Alcalá. http://preseea.linguas.net Consulted: 2015-Present.

Rivers, Wilga M. (1992) The Undergraduate Program 1: Courses for All Comers. *Teaching Languages in College.* Lincolnwood, Illinois: National Textbook Company.

Savignon, Sandra J. (2000) Communicative language teaching. En Michael Byram (Ed.) *Routledge Encyclopedia of Language Teaching and Learning.* (pp. 125–129). London: Routledge.

Stern, H.H. (1992) *Issues and Options in Language Teaching*. (pp. 132) En Patrick Allen and Birgit Harley (Ed.) Oxford: Oxford University Press.

Tedick, Diane J., (Ed). (1997) Proficiency-Oriented Language Instruction and Assessment: A Curriculum Handbook for Teachers. *The Minnesota Articulation Project*. Minneapolis: Center for Advanced Research on Language Acquisition, University of Minnesota.

Tschirner, Erwin. (1996) Scope and Sequence: Rethinking Beginning Foreign Language Instruction. *The Modern Language Journal 80*, (pp. 1-14).

APROXIMACIÓN A LA CONQUISTA ARABO-MUSULMANA DE LA PENÍNSULA IBÉRICA EN 711 A TRAVÉS DE LA TOPONIMIA.

APPROACH TO THE ARAB-MUSLIM CONQUEST OF THE IBERIAN PENINSULA IN 711 THROUGH TOPONYMY.

José Beneroso Santos
Doctorando en Historia
por la Universidad Nacional de Educación a Distancia

Recibido: 18 de septiembre 2021. Aceptado: 1 de diciembre 2021

Resumen: La invasión y conquista arabo-musulmana de Hispania a partir de 711 es uno de los hechos históricos más importantes acaecidos en la península Ibérica. Entre los aspectos menos conocidos de este proceso se encuentra el lingüístico. El estudio toponímico ha pasado a ser un recurso imprescindible para avanzar y profundizar en el conocimiento de los momentos iniciales de al-Andalus.

Palabres clave: Al-Andalus, *al Yazira al-Jadrá*, Gibraltar, *al-Buhaira* y *Wadi Lakko*

Abstract: The arab-muslim invasión and conquest of Hispania starting in 711 is one of the most important historical events that occurred in the Iberian Peninsula. Among the lesser known aspects of this process is the linguistic one. The toponymic study has become an essential resource to advance and deepen the knowlegde of the initial moments of al-Andalus

Keywords: Al-Andalus, *al Yazira al-Jadrá*, Gibraltar, *al-Buhaira* y *Wadi Lakko.*

A MODO DE INTRODUCCIÓN

La invasión y conquista arabo-musulmana de Hispania a partir de 711 es uno de los hechos históricos más importante acaecidos en la península Ibérica. La zona norte del estrecho de Gibraltar alcanzó pronto un gran protagonismo puesto que estas tierras fueron testigo de la incursión de Tarif ibn Malik en 710, preludio de lo que al año siguiente, en 711, sería ya la invasión dirigida por Tariq ibn Ziyad al frente de un ejército mayoritariamente integrado por bereberes de la antigua *Mauretania Tingitana*, y de la campaña «oficial» de Musa ibn Nusayr en 712, en la que aparece un gran número de árabes y bereberes de la zona de la *Mauretania Caesariensis*.

Estos hechos hay que enmarcarlos en la profunda crisis sufrida por el mundo mediterráneo entre finales del siglo VII y principios del siglo VIII, que afecta de un modo u otro a todas las sociedades estatales y tribales existentes. En este sentido, la ocupación y dominio de lo que será al-Andalus es consecuencia directa de la política expansiva que venía desarrollando el califato de Damasco, en la denominada «segunda ola de conquistas».

La situación político-social en Hispania era inestable y la irrupción arabo-bereber sirvió de catalizador en el proceso de desestructuración por el que atravesaba el reino visigodo a primeros de siglo VIII, provocando un viraje en los acontecimientos peninsulares, pasando de unas formas latino-cristianas a otras arabo-islámicas. Es por lo tanto un proceso que no sólo afecta a los planos político y social sino también al económico, al religioso y al cultural.

Un aspecto a resaltar en esta conquista es la transición lingüística que se produce en el Mediterráneo occidental desde la llegada de los grupos arabo-musulmanes al norte de África hasta la culminación con la conquista de Hispania. Tal contingencia sustituye el latín por el árabe en el ámbito político-cultural y favorece la aparición de otras formas dialectales. Este cambio debe contemplarse dentro del proceso

evolutivo del latín en el Mediterráneo africano, más concretamente en la zona del Estrecho de Gibraltar, centro gravitatorio de la conquista musulmana de la Península Ibérica, y afecta a ambas orillas, desarrollándose desde mediados del siglo V hasta prácticamente la segunda mitad del siglo X, con la consolidación del califato de Córdoba.

Una de las consecuencias de la llegada de los arabo-bereberes en el 711, la que más nos interesa en esta ocasión, fue el rápido cambio operado en la toponimia del cono sur peninsular, particularmente en la zona más próxima a actual la bahía de Algeciras, zona de gran relevancia durante las operaciones militares iniciales.

DE HISPANIA A AL-ANDALUS O DEL LATÍN AL ÁRABE PASANDO POR EL ROMANANDALUSÍ

Uno de los aspectos más significativos en el estudio del proceso conquistador musulmán en la península Ibérica es el de la lengua utilizada por los hombres que acompañaron a Tariq ibn Ziyad en 711.

Asimismo, para el estudio de al-Andalus es esencial conocer la situación del latín, tanto en Hispania como en las provincias romanas africanas, después de las incursiones de los pueblos germanos, en particular visigodos y vándalos.

El latín practicado en tierras africanas evolucionó progresivamente hacia unas formas dialectales afrorrománicas[1], que fueron las encontradas por los árabes a su llegada y las utilizadas para comunicarse. Luego, y aun teniendo en cuenta que en la zona del Mediterráneo occidental se produjo un cambio lingüístico a partir del siglo VI[2], podemos afirmar que todavía a finales del siglo VII y principios del VIII el latín era la lengua de comunicación por excelencia de las poblaciones existentes en ambas orillas del Estrecho. En realidad, conocemos que en *Septem* y *Tingi* se utilizaron variantes del latín a principios del siglo VIII, como una serie de inscripciones así lo pueden confirman.

Es evidente que en el norte de África, tal como señala Villaverde, «el latín coexistía con otras lenguas especialmente, diversas formas dialectales *tamazight* [bereber], hebreo y griego, lo cual confirma la complejidad multicultural del país en época tardorromana»[3], en referencia a la *Mauretania Tingitana,* pero de igual modo aplicable a la *Mauretania Caesariensis*, y nada hace pensar que esto no continuase siendo así, debido fundamentalmente a la «difusión y arraigo del cristianismo»[4], aunque quizá ya no con la misma intensidad.

Recordemos que el modelo cultural latino-cristiano estaba fuertemente enraizado en estas tierras antes del dominio musulmán, por la presencia romana y bizantina (esta vigorizó el uso del latín en detrimento del griego utilizado en Oriente), y por la presencia de visigodos y vándalos[5], que profesaron el arrianismo, y que también utilizaron el latín como lengua vehicular, pues apenas hablaban ya germánico.

Como consecuencia el cristianismo favoreció la consolidación y perpetuación de la latinización de esta zona, que fue mucho más intensa de lo que habitualmente se ha venido aceptando, y que todavía era visible a finales del siglo VII. En este sentido, el cristianismo fue un factor de aglutinación de los pueblos norteafricanos y a su vez sirvió de nexo con la población de la Bética.

A grandes rasgos y de forma generalizada defendemos que se siguió utilizando el latín y varios dialectos afrorrománicos en el ámbito urbano; el bereber en el ámbito tribal, en las montañas y zonas desérticas y probablemente pudieron ser utilizadas indistintamente ambas lenguas en el medio rural y sedentario. Desde finales del siglo VII, el árabe se fue imponiendo de forma lenta.

Hablamos de una población, tal como señala entre otros Ibn Jaldún, creemos que haciendo una clasificación etno-lingüística (en la que a la vez se ve reflejada perfectamente los factores sociales, económicos y culturales), que estaba compuesta esencialmente por bizantinos, los *rum*; latifundistas romanizados, los *afrang*; habitantes de los principales núcleos urbanos, muchos de ellos comerciantes y artesanos,

en su mayoría bastante cristianizados, los *afariqa*; y los autóctonos, que presentaban estructuras clánico-tribales en las que el factor agnático era fundamental, no culturizados y muy belicosos, los *barbar*, habitantes de las montañas y del medio rural, éstos últimos sujetos a prácticas sedentarias.

Por lo tanto, gran parte de los hombres que llegaron con Tariq ibn Ziyad en el 711 empleaba como lengua vehicular una forma dialectal del latín. Para el profesor Marcos-Marín estos guerreros eran: «hablantes afrorrománicos, moros [bereberes], que llegaron en gran número a la península Ibérica a partir del 711. Por razones lingüísticas evidentes, sus hablas entraron en contacto con el iberrománico y algún papel tuvieron que representar en su evolución, particularmente la del romance andalusí, pero no sólo de éste»[6]. Es cierto que todavía en la actualidad desconocemos con certeza y en gran medida muchos aspectos sobre la lengua o lenguas utilizadas por estos conquistadores, en su mayoría bereberes, entre los que aparecían muchos *mawali*. No obstante, hay que destacar que es evidente el contacto entre dialectos afrorrománicos, aún pocos conocidos, e iberorrománico[7], lo que sin duda facilitó la conquista[8] y el posterior dominio arabo-bereber sobre el territorio hispano. Un contacto que explicaría la rapidez en la evolución de los acontecimientos y muchas de las condiciones en las que se produjo la aniquilación del reino visigodo y el surgimiento de al-Andalus.

En realidad, el tránsito de personas entre el norte de África y la península Ibérica ha sido, prácticamente, una constante prehistórica y nunca se vio afectado, aunque es cierto que hubo periodos con una mayor o menor afluencia. Así, durante época romana el flujo se mantuvo y el intercambio lingüístico tampoco se interrumpió, puesto que a pesar de que gran parte de los territorios ocupados por los imperiales fueron recuperados por los visigodos en 624, la zona del Estrecho de Gibraltar, con los núcleos poblacionales de *Septem*, *Traducta* y *Asidona*, siguió funcionando como una unidad socio-mercantil, sin tener en cuenta el cambio político-militar producido.

Muy interesante también es lo señalado por Karima Bouallal que dice: «El latín no desapareció bruscamente, a pesar de la llegada de los árabes, porque ellos mismos lo utilizaron bastante tiempo; el tiempo necesario para los procesos de islamización y arabización. Los árabes, ante la presencia de dos idiomas —el bereber (lengua local oral) y el latín (lengua de prestigio escrita)—, mantuvieron el uso del latín. El abandono del latín no podía ser inmediato porque bastantes comunidades bereberes [profesaban] el cristianismo. Entonces el único medio de hacer llegar la nueva verdad a los autóctonos era el uso del latín»[9].

Con respecto a esto, recordemos que las fuentes hablan de una presencia árabe simbólica en las tropas de Tariq Ibn Ziyad que inician la conquista y que estuvieron asentadas en Tánger y que Julián era *sayj* de los *ghumara*, una tribu bereber cristiana y muy romanizada; de hecho, uno de sus generales era Mugit *al-rumí* (el cristiano).

Para Chalmeta: «Musa había dejado cuantos bereberes (auxiliares y rehenes) le acompañaban, bajo el mando de Tariq, en Tánger, convertida en ciudad-cuartel [*misr*] Su número se elevaba a 12-19000 hombres, a los que hay que añadir unos pocos [...] árabes. La cifra que las fuentes asignan a estos últimos oscila entre 12 y 17; cantidad suficiente para su cometido de "enseñar el *Qu'ran* y las (normas) islámicas a los bereberes". [...] Llama inmediatamente la atención [...] que los seguidores de Tariq son esencialmente no-árabes: bereberes y *mawali*. Todas las fuentes dan el elemento árabe como una ínfima minoría [...] y parecen estar aludiendo a una o dos decenas»[10].

Por lo tanto, el proceso de islamización y sobre todo de arabización, estaba todavía sin realizar en el Magreb cuando se inicia la invasión de Hispania. No hubo tiempo material con un número tan reducido de árabes para instruir a miles de bereberes, que además desconocían el árabe. Ciertamente el proceso se aceleró en los años siguientes una vez llegados los contingentes militares a la Península.

Aun así, y pasados algunos años, este proceso de islami-

zación y de arabización tampoco se desarrolló en la zona del Magreb de forma simultánea, apareciendo distintos ritmos en su evolución. De igual modo sucedió en al-Andalus. En ambos territorios la arabización se desarrolló más pronto y con mayor intensidad en el medio urbano. Por otra parte, es difícil de entender este proceso de arabización, en particular en lo referente al cambio lingüístico, sin tener en cuenta el contexto social en que tuvo lugar.

Si acaso, incluso sin completarse la islamización de la población norteafricana, sí aparece desde los preludios de la invasión el llamamiento a la *yihad*, que actúa como móvil ideológico, tal como figura en los *fulus* acuñados para financiar la campaña y es señalado por Guichard[11]. Estas monedas con leyendas latinas siguen siendo emitidas y utilizadas en Ifriqiya hasta bien entrado el siglo VIII, lo que puede confirmar que el latín era la lengua usada por la mayor parte de las tropas. Desde esta perspectiva, son la religión y, en menor medida, la lengua y la cultura árabes las que determinan y conforman la columna vertebral del proceso expansivo musulmán, en definitiva, de la civilización arabo-musulmana que quedará establecida a ambos lados del Estrecho.

Además, en nuestra opinión, el bilingüismo de las monedas estaría justificado porque estas piezas, con marcado carácter local, de acuñación rápida y de emisión más reducida, no solo estaban destinadas para los arabo-bereberes, es decir, para gente que dominaba aunque fuese mínimamente el árabe, sino también para gran parte de las tropas del *comes Iulianus*, entre los que se encontraban *ghumaríes* no islamizados (que no hablaban árabe ni entendían dicha grafía), y para sufragar los gastos logísticos de la operación: barcos, marinería, impedimentos, etc.

Este hecho es sumamente importante, es decir, la acuñación de los primeros *fulus* con leyenda en latín y árabe en los momentos anteriores a la invasión (puesto que estas monedas fueron las utilizadas para pagar las soldadas en la conquista de *Tingi* en 709), nos está indicando la existencia de un bilingüismo real entre las tropas que intervienen en la

conquista. Son monedas en las que también se alude a la *yihad* y estarán en circulación, debido a la gran cantidad emitida, hasta algunos años después de producirse la conquista de la Península.

Por otro lado, llama la atención la escasez de términos bereberes en el árabe andalusí. A la espera de algunas investigaciones ya iniciadas por el profesor Corriente y que han tenido continuidad, debemos señalar que el bereber utilizaría preferentemente para comunicarse en Hispania algunas formas afrorrománicas que simplificarían la comunicación con la población autóctona y que sí verterían algunos términos al romandalusí.

Además, sospechamos, a la espera de ser confirmado por un necesario estudio aún pendiente de realizar, que los préstamos bereberes tuvieron que verse pronto reflejados en la toponimia del sur peninsular. El ejemplo de *Tsur Arab*, muralla de los árabes, que aparece en las fuentes como primer enclave en Gibraltar puede ser uno de ellos. La voz *tsur* es bereber y significa construcción fortificada, recinto resistente, y con un matiz de aparecer aislado. Un término, a su vez, que según Volney, proviene del fenicio con el significado de roca y lugar fuerte[12].

Por último, hay que destacar que una población de aproximadamente cuatro millones de personas como tenía en esas fechas Hispania no pudo incorporar el árabe de forma rápida y eficiente. Una lengua que en el mejor de los casos era hablada por unos treinta y cinco mil guerreros conquistadores, máxime cuando la mayoría de estos eran bereberes procedentes de las antiguas provincias romanas del norte de África, como acabamos de señalar. Unos territorios que habían sido intensamente romanizados y cristianizados y que habían ofrecido una fuerte resistencia al dominio musulmán, basta recordar los episodios protagonizados por la *Kahina*, al parecer sacerdotisa y *sayj* tribal, y que conservaban otras lenguas para comunicarse.

ALGUNOS APUNTES SOBRE EL ROMANDALUSÍ, EL TODAVÍA FRECUENTEMENTE MAL LLAMADO MOZÁRABE.

A partir del siglo VIII, ciñéndonos a la península Ibérica, la arabización supuso sin duda una deslatinización, pasando, a grandes rasgos en términos lingüísticos, la población indígena del latín, fundamentalmente en sus formas dialectales iberrománicas y afrorrománicas, al árabe a través del romandalusí. Lo que conllevó la aparición y adaptación de palabras que, aun manteniendo en muchos casos una forma, o parte, latina, son modificadas con elementos árabes, dando lugar a una gran cantidad de términos romandalusíes visibles en los textos musulmanes conservados, y que tanta importancia adquieren para el conocimiento de al-Andalus, particularmente en su época inicial.

Resulta evidente que los sistemas fonémicos del árabe, del bajo latín peninsular y africano, y también, aunque en menor medida, del bereber, «que entran en contacto a raíz de la conquista islámica de la Península Ibérica [dando lugar al romanandalusí] eran considerablemente dispares»[13] y, también lo es que aunque además «la primera lengua [el árabe] ocupa de resultas una situación dominante en la sociedad del nuevo país emergente, Alandalús, práctica y progresivamente bilingüe casi desde sus comienzos y hasta culminar la tendencia al monolingüismo árabe hacia fines del siglo XII [...], la lengua dominada, el haz dialectal romandalusí se mantiene con considerable aunque decreciente vigor en el uso doméstico, rural y jergal, por razones demográficas y prácticas, hasta el siglo XI, en que, más por prurito de demostrar [*arabidad*] y superioridad sobre los aborrecidos aliados y ocupantes bereberes que por distanciarse lingüísticamente del enemigo cristiano, se acentúa y completa en pocas décadas su proceso de eliminación. Durante esos aproximadamente cinco siglos de contacto estrecho entre árabe y romance tienen lugar tres procesos que afectan decisivamente a la fonología de los arabismos, a saber, a) la emergencia del haz dialectal andalusí,

con ciertos rasgos fonémicos y de otros niveles resultantes de la interferencia del romance, con el que coexiste en situación de *Sprachbund*, b) la evolución del haz dialectal romandalusí, con recíproco condicionamiento por parte del árabe, y c) la aparición de los primeros arabismos de los romance septentrionales, procedentes directamente del andalusí o por mediación del romandalusí, que también hace préstamos propios al andalusí y a los romance septentrionales»[14].

Las diferencias lingüísticas entre el árabe y el latín, básicamente en sus formas dialectales del iberromance y el afrorromance, eran enormes, por lo que las influencias quedaron en un primer plano, en los aspectos más superficiales, sobre todo y particularmente en el léxico. Esto hay que tenerlo presente porque supuso la aparición de nuevas formas. Formas, híbridas o mixtas, en su mayoría en el ámbito médico, en el botánico, en el comercial, en el agrícola, en el doméstico…, y lo que para este trabajo más nos interesa, en el toponímico. En este ámbito en particular la influencia árabe es bastante significativa, y la bereber parece que también, aunque ha sido menos estudiada.

Sin profundizar en el asunto, y centrando la atención en el plano morfológico, debemos señalar en concreto la anteposición del artículo al- a numerosos sustantivos que dieron lugar a una gran cantidad de topónimos en la zona estudiada. La fusión de este artículo con el lexema del sustantivo permite el uso de artículos romances (casos de la *almoraima*, la *aljaranda*, el *alqantir*…). En muchos otros casos las palabras se han formado con un artículo árabe como al- unido a una base latina, como ocurre en los casos de alpiste <árabe *al*+latín *pistu*, (siendo un ejemplo el topónimo menor: Marjal del alpiste) y almonte <árabe *al*+latín *monte*, o entre palabras o formas distintas, dando lugar a una gran cantidad de términos que abundan en los primeros tiempos de al-Andalus, como pueden ser los casos de Fontetar <latín *fonte*+árabe *thar*; Guadacorte <árabe *Wadi*+ latín *cohorte* (o *cortex*), etc. Acerquémonos, insistimos, sin entrar mucho en la materia, porque superaría con creces las pretensiones de este trabajo,

a varias cuestiones lingüísticas que aparecen desde los primeros momentos de al-Andalus, que por su importancia y repercusión deben ser expuestas.

Los sistemas fonológicos que entran en contacto, principalmente a partir de 711, el haz dialectal árabe, el bereber[15], el iberromance, el afrorromance y, posteriormente, el romandalusí van «a ajustarse a un cuadro de equivalencias regulares, aunque diacrónica y diatópicamente variables también regularmente»[16], provocando una serie de correspondencias, sustituciones y combinaciones entre ellas, tanto en el plano fonético como en el morfológico-sintáctico. Además, se debe aceptar una vigencia plena del romadalusí hasta la instauración del Califato cordobés y la política omeya de *Umma Wahidum*.

Por otro lado, nos parece conveniente señalar lo siguiente: los árabes utilizaron la «c» del romance con que reproducían la «c» latina ante las vocales «e» e «i». Fundamentalmente por esta razón se produce el predominio de la «c» en las reproducciones de palabras romances y el empleo abundante de «ch» por «c» en topónimos de al-Andalus, de los que la zona estudiada presenta algunos casos como *Marsa al-Cheyera*.

Una peculiaridad del andalusí es la imela o imala, es decir, el paso de «a» a «e» y después a «i» (a-e-i), que explica un gran número de topónimos[17], como ocurre con *Hispalis >Isbilia>Sevilla* y que se dio con bastante frecuencia desde los comienzos de al-Andalus.

También es de destacar que: «El paso de palabras árabes, tomadas del registro oral en la mayoría de los casos, al romance exigió su reajuste fonológico y acomodación al sistema hispano, pues muchos de los fonemas árabes no tenían correspondencia en español»[18].

En este proceso aparecen de forma manifiesta como rasgos fonológicos destacados el vocalismo y el consonantismo. Así, «el sistema vocálico [...] árabe, con tres fonemas/a, i, u/, dos grados de abertura y distinción de cantidad se acomodó al del romance, con tres grados de abertura y perdiendo la función distintiva de cantidad»[19].

Sin duda el vocalismo establece grandes diferencias entre los sistemas fónicos árabe y romance porque, y esto es preciso resaltarlo, « el primero exhibe un sistema triangular escueto (/a/, /i/ y /u/), sin fonemas intermedios, mientras que el segundo tiene uno o dos, según fechas y áreas, entre las vocales cerradas y la abierta, lo que supone un sistema de cinco (/a/, /e/, /i/, /o/ y /u/), o siete vocales […]»[20], aunque es cierto que el contacto entre ambos no se ve reflejado en un incremento de los fonemas vocálicos del andalusí.

Además, y al mismo tiempo, se produce la adaptación de fonemas consonánticos árabes siguiendo las mismas pautas evolutivas de los fonemas romances. Todo esto provocó una gran amalgama de «nuevos» términos que, particularmente en el campo de la toponimia, consideramos de gran importancia para este estudio.

ALGUNOS TOPÓNIMOS PARA UNA POSIBLE VERTEBRACIÓN DE LOS HECHOS

Un recurso al que se suele acudir en busca de información en el estudio de la entrada de los arabo-bereberes en la península Ibérica, en particular desde hace una década, es el registro toponímico. Pese a los errores que ha inducido con frecuencia, consideramos la toponimia un registro imprescindible para realizar estudios en el ámbito arqueológico e histórico, puesto que proporciona información fosilizada que permite completar la ya extraída de otras fuentes. Y hasta es bastante frecuente que los topónimos digan lo que omiten o callan los documentos.

Los hechos que tratamos son de sobra conocidos. Siguiendo un eje cronológico-factual de los acontecimientos acaecidos en 711 con la entrada de los arabo-bereberes en la Península, y apoyándonos en una serie de topónimos, hemos intentado vertebrar lo sucedido recurriendo a los siguientes topónimos: al-Andalus, península Ibérica, tierra de *yihad*; Gibraltar, punto de referencia geográfico en la invasión;

al-Yazirat al-Jadrá, primera ciudad fundada por los musulmanes; *al-Buhayra*, principal escenario de los acontecimientos bélicos y *Wadi Lakko*, lugar del gran enfrentamiento entre las tropas de Tariq ibn Ziyad y Rodrigo

Al-Andalus

El nombre de al-Andalus referido a la península Ibérica aparece documentado por primera vez en unas monedas fechadas pocos años después de la llegada de los arabo-musulmanes, concretamente en 716. Es una emisión bilingüe, ya referida, en la que en una cara de las monedas figura la leyenda en latín «Span», de *Spania*, y en la otra en árabe «al-Andalus». No obstante, pese a la clara identificación de este topónimo con la Península como sinónimo de *Spania* o Hispania, no existe unanimidad en cuanto a su origen.

Se han ofrecido varias respuestas sobre el origen de este término, destacando la conocida ya en época medieval y difundida a partir del siglo XVI como tesis vándala, argumentada por el arabista Reinhart Dozy siguiendo criterios filológicos. Dozy señalaba que el actual nombre de Andalucía derivaba del etnónimo *Vandalicia*, tierra de los vándalos, en referencia a la Bética; pero sabemos que aquel término es una castellanización, bastante conocida por cierto, del adjetivo árabe *Al-Andalusiyya*, derivado del sustantivo al-Andalus, y que se caería en un grave error si se invierte su posible derivación. De ese modo *Vandalicia*, o *Vandalia*, que significa «la tierra de los vándalos», haría clara alusión a lo que conocemos actualmente como Andalucía, atribución muy en boga durante el siglo XIX y primera mitad del XX.

Sin embargo, Corriente indica que «Alandalús[21]<jazíratu lPandalus, probable corrupción del griego (*he nesos tes*) atlantídos "isla de Atlantis", recibido directamente o a través del siriaco gazarta d+atlantidos, como parte del legado de los mitos platónicos que se difundieron con la filosofía y hasta con la mera cultura griega [...] favorecido por la etimología

popular que introducía así la raíz del árabe dalas "tiniebla" en lo que para los orientales era su región y mar característicos»[22].

Asimismo, el profesor Vallvé también identifica el término Al-Andalus con Atlántida: «Desde un punto de vista fonético, se puede explicar fácilmente el paso de Atlas o Atlantes a al-Andalus. De Atlantis pudo derivar Al-lantis por asimilación t por l, como de Atlético decimos al-lético, y a través de Al-lantis se puede explicar la Yazira al-Andlis o Yazira Al-Andalus, nombre dado a la Península por los árabes»[23]. Tampoco descarta que su origen, siguiendo a algunos geógrafos árabes, como Ibn Rustah, pueda provenir de la identificación con Atlántico o Mar de las Tinieblas[24]. Por otra parte, el mismo profesor señala, algo que nos parece sumamente interesante, que Ibn Atir (siglos XII-XIII) dice que el emperador Heraclio (siglo VII) habla de la recaudación de unos impuestos de las tierras de Egipto, Ifriqiya y al-Andalus[25].

En nuestra opinión, el término deriva de la expresión germana *land hlauts*. Seguimos entre otros autores a Halm, que se inclina por una traducción goda de la *habita sorte* tardorromana, concedida después de un período de lucha constante en los primeros años del siglo V (concretamente entre 409 y 417)[26], y contemplada como *land hlauts*, «tierra repartida en lotes», tierra de sorteo, lotes, reparto, etc., a pesar de que este término aparece generalmente como *gothica sors* en las fuentes clásicas en latín para referirse a las tierras del *Regnum Gothorum*. No obstante, en lengua germánica bien podría corresponderse con el término *landhlauts*, (de *land,* tierra, y *hlauts*, lote, pero en alusión al fraccionamiento de un todo otorgado por sorteo).

Es manifiesta la relación de tipo histórico que puede haber con los visigodos, puesto que «en los territorios fronterizos, desde el reinado de Wamba, habían sido instalados una serie de asentamientos[27], fortificados, con colonos militares a los que les fueron concedidos lotes de tierras al igual que se había procedido anteriormente con el reparto de las *sortes*[28], que se encargarían de la defensa ante un eventual ataque

del exterior»[29]. El origen de estas *sortes* hay que buscarlo en el *ius hospitalitatis* romano, con el que los soldados visigodos quedaban establecidos de forma permanente, concediéndoseles tierras, las *sortes gothicae*.

La evolución habría sido *Landhlauts>Landalos>Al-Andalus,* desde el punto de vista fonético perfectamente aceptable; sin embargo, fue llamada así por los habitantes del norte de África, reconociendo en «Landalos» *una voz afrorrománica y*, posteriormente, con la arabización, el término pasó a «al-Andalus». Y porque además para los grupos arabo-musulmanes esta tierra de *yihad* era, principalmente, una tierra de reparto de botín, y el acceso a la tierra se realizó habitualmente y de forma mayoritaria por lotes concedidos a grupos siguiendo un modelo clánico-tribal. Algo a lo que aspiraron los bereberes y los *afariqa* que acompañaron a Tariq ibn Ziyad «el Tuerto» en la conquista de Hispania. Esta razón, entre otras, hace pensar que no debemos descartar tampoco un posible origen *amazigh* (bereber) del término Al-Andalus, pero no amparado en la tesis que defiende la expresión bereber *tamort uandalos* (tierra de los vándalos), pues tampoco existen pruebas documentales históricas ni registros arqueológicos que la avalen. Quizá sí sería posible como una deformación del término híbrido (bereber-godo) y contraído de *tamort landalos* (tierras de reparto o tierras repartida en lotes).

La opción de hacer derivar al-Andalus del término *land hlauts* a través de una voz afrorrománica evitaría un importante contratiempo existente en la actualidad para aceptar plena y definitivamente dicha tesis; los vocablos godos que pasaron al romandalusí son muy escasos e inexistentes en el árabe[30].

Gibraltar

La imbricación de términos orográficos y antropomórficos es muy abundante en la toponimia peninsular. Un claro ejemplo es Gibraltar, *Yebal Tariq*, Montaña de Tariq, ate-

niéndonos al significado habitualmente más aceptado. Cabe la posibilidad también de que su origen sea *Yebal Tarif*, en referencia a otro protagonista de los sucesos de 710-711.

Independientemente, Tarif puede provenir de la forma *Tarf*, en el sentido de «el que abre el camino» o «el primero», o de *tarf*, acantilado, precipicio, punta de tierra…, de la que derivaría Trafalgar (*Tarf al-garb*, «punta de occidente»). En este caso Gibraltar derivaría de *Yebal al-tarf*, montaña del acantilado. De hecho, existe un topónimo aquí, Los Tarfes, citado por Hernández del Portillo en el siglo XVI[31].

Además, también se contempla una posible derivación de la forma *Yebal al-fath*, Montaña de la victoria, quizá como consecuencia de la ciudad almohade fundada en este lugar, *Madina al fath*, Ciudad de la victoria, e incluso, como señala Recio[32], *Gibeltah* o Montaña de la entrada.

Así mismo, las fuentes cristianas denominan a Gibraltar *Promonturiis Traductinis*, en clara referencia al núcleo poblacional más importante que se hallaba al otro lado de la bahía, *Iulia Traducta*, crecida en detrimento de la antigua *Carteia*. Las fuentes árabes denominaron a *Traducta*, o más exactamente a una población *ex novo* de tipo *misr* (ciudad-campamento) que surge a su lado, *Al-Yazirat al-Jadrá*, para nosotros probablemente en referencia al peñón de Gibraltar, como a continuación veremos.

Al-Yazirat al-Jadrá

Es posible que, en los primeros momentos de la invasión, cuando los musulmanes se refieren a Gibraltar la llamasen *al-Yazirat al-Jadrá*, la isla o península verde, por el aspecto destacado que ofrecía en cuanto a su vegetación, y en comparación con los islotes y cabos de las proximidades en esta área del Estrecho totalmente exentos de vegetación, y principalmente áridos y rocosos. Poco después se designó con este nombre una nueva población, como hemos referido, junto a la *Traducta* romana. Era en esos momentos el núcleo más

relevante que encontraron a su llegada desde África por la ruta (*mayaz* o paso) más conocida y empleada, aunque no la más fácil.

También se hace coincidir este término con un islote existente hasta principios del siglo XX, hoy integrado en las estructuras portuarias de la ciudad de Algeciras, denominado Isla Verde. Asimismo, se ha identificado *Al-Yazirat al-Jadrá* con la península Ibérica, quedando fosilizado el topónimo en la ciudad de Algeciras, insistimos, por ser la población de mayor importancia que encuentran los musulmanes a su llegada desde África por la ruta más frecuentada. De igual modo se denominó *al-Yazirat Qartayyana* a *Carteia* y *al-Yazirat Tarif* a Tarifa, e incluso, en algunos casos, de forma puntual también a esta última *al-Yazirat al-Andalus* o *al-Yazirat-al-Jadrá*.

Al-Yazirat Jadrá tomó pronto particular relevancia al convertirse en la verdadera retaguardia del ejército arabo-bereber, erigiéndose así en el principal centro de las operaciones militares y cabeza de puente para el desembarco del resto de las tropas arabo-musulmanas que acompañaban a Tariq, y por quedar instalado en este punto un hospital de campaña.

Al-Buhaira

Entre las fuentes más antiguas, como es el caso del *Ajbar Maymua*, aparece el término *al-buhaira* para señalar la batalla o el encuentro del «río del lago». En concreto señala: «Encontráronse Rodrigo y Tarik, que había permanecido en Algeciras, en un lugar llamado el Lago, y pelearon encarnizadamente [...]»[33].

Consideramos que cuando se menciona *al-buhaira*[34] no solo se hace referencia a «lago», «albufera», «marisma»... sino que puede ofrecer además otro significado no menos importante como es el de «mar pequeño», e incluso el de «bahía». Así este término adquiriría un mayor sentido, tal como aparece en la obra de Al-Maqqari: «[...] escribió a Musa para pe-

dir ayuda, diciéndole que había tomado Algeciras, un puerto de [al-] Andalus, quedando en su posesión el paso a este país; que había sometido su distrito hasta la bahía; pero que Roderico estaba ahora avanzando contra él con una fuerza que no estaba en su poder resistir»[35].

También el *Ajbar Maymua* dice lo siguiente: «Apenas llegó esto [envío de un gran ejército visigodo] a noticia de Tarik, escribió a Musa pidiéndole más tropas y dándole parte de que se había hecho dueño de Algeciras y del lago»[36]. De nuevo insistimos, aquí «lago» cobraría un mayor sentido si se interpreta como «bahía», concretamente como la zona que bordea la bahía, el arco de la bahía, es decir *Yebal Tariq* (atalayas o fortificación), *Carteia,* los restos de *Portus Albus* y algunos otros asentamientos menos relevantes y, por supuesto, la propia *Iulia Traducta.*

Por lo tanto, *al-buhaira* o *al-buhayra* es una forma diminutiva del étimo *bahr*, mar, por lo que la posibilidad de identificar el lago de los textos con la actual bahía de Algeciras, en definitiva, un «pequeño mar, mar menor o marecito», no debe ser desechada, de tal manera que el río del lago—*Wadi al-buhaira* o *Wadi lakko*—de las fuentes, como enseguida analizaremos, se correspondería, sin duda, con el Guadarranque, su principal río y el término «lago» con dicha bahía.

Wadi-Lakko

El primero que recoge el término *Wadi Lakko* es Ibn Qutiyya en el siglo X, y lo hace escenario de un enfrentamiento entre Tariq ibn Ziyad y Rodrigo. Descendiente de godos, debía emplear habitualmente términos en la lengua nativa de sus ancestros, utilizando muchas veces palabras híbridas. En este caso en concreto, cambió el término árabe *al-buhayra*, al que acabamos de referirnos, por el románico *lakko*, para referirse a lago o bahía.

Además, es uno de los casos más claros de puesta en escena del paso de la transmisión oral a escrita llevada a cabo

por los discípulos. En estos escritos aparecen formas iberorromances o romandalusíes mezcladas con árabes, que mostrarán unas características propias. En este caso *Wadilakko* <árabe *wadi*+ iberorromance *lacco*.

En cualquier caso, se trata de una denominación genérica o perífrasis, *Wadi Lakko*, *Wadi l-buhayra*, *Wadi al-bahr*..., que podría aplicarse perfectamente al actual río Guadarranque, o a su valle. Al-Himyari (siglo XV), basándose en la información del geógrafo al-Bakri (segunda mitad del siglo XI), quien, al describir *Qartayyana al-Yazirat*, es decir *Carteia*, llama a su río, el Guadarranque, *Wadi l-bahr*, es decir «río del mar»[37] y también en algún momento *Wadi l-ramal*, río de la arena[38].

Muy interesante es lo señalado por Ibn Hayyan (siglos X-XI), citado por Al-Maqqari (siglos XVI-XVII), que habla del «río del lago» de la tierra de Algeciras (*Wadi Lakko min ard al-Yazirat al-Jadrá*), «de la costa frontera [pensamos que con la intención de diferenciarlo de otro *Wadi Lakko* existente], al lugar de paso»[39], por donde principalmente entraban los musulmanes, esto es el *majaz*, paso, de la bahía[40]. Por lo que ambos localizan el enfrentamiento en las inmediaciones de ésta.

Sánchez-Albornoz, siguiendo a Ximénes de Rada en su obra *De Rebus Hispaniae*, identifica y defiende con encono *Wadi Lakko*, intencionadamente convertido en *Wadi Lakka*, con el Guadalete (que la citada obra transcribe como Vadalec, Guadalec o Guadalet) como escenario de la batalla. Argumenta que *lakko* es un error de un copista de Al-Himyari que tenía que haber traducido *lakka*, en referencia a una antigua población romana conocida como *Lacca*, que localiza en el curso medio del actual Guadalete, y de la que apenas existe información fehaciente[41]. Para nosotros este *Wadilakko* no se corresponde, en ningún caso por deformación o evolución toponímica, con el Guadalete, y sí, geográficamente, con el Guadarranque.

Por último, señalaremos que, aunque el hidrónimo Guadarranque se ha hecho derivar durante años de *Wadi ar-ramk*,

río de las yeguas, incluso de *Wadi al-Rinq*, según Vallvé en referencia al rey Rodrigo[42], pensamos que su etimología es mucho más compleja a pesar de que bien pudo derivar del término árabe o bereber *rmel*, arena, dando *Wadi ar-rmel*, río de la arena. Este término es muy sugestivo, puesto que el principal problema que tuvo *Carteia* para mantener la actividad portuaria fue la colmatación de su río, que provocó la inutilización de las instalaciones portuarias y la decadencia como ciudad en el siglo V.

Sin embargo, para nosotros, el origen es otro bien distinto. Podemos establecer que los «lagos de pesca» referidos en el *Digesto*, además de los caladeros en sí, contemplaban también el puerto, las instalaciones y la ensenada donde se hallaban. En concreto, la consideración que se le hace a *Carteia* dentro de los «*lacus quoque piscatorios et portus in censum dominus debet deferre*»[43], es definitiva para el uso del término «lacus» en esta zona. Es decir, este lugar generaba pesquerías cuyo aprovechamiento era necesario declarar en la *forma censualis*, tal como señala el *Digesto*[44].

No es descabellado, por ende, pensar que la denominación, con un claro matiz técnico, en época anterior a la llegada de los arabo-bereberes fuese *Fluvium lacus quoque piscatorios et portus in censum Carteiensis*, porque *Carteia* debió conservar durante todo el Bajo Imperio y época visigoda cierta actividad pesquera, siendo el Guadarranque la principal vía de comunicación y salida hacia el exterior de sus producciones. Para los musulmanes pronto sería el «río de la arena del lago», *Wadi ar-rmel al-buhayra,* haciendo así distinción del otro *Wadi ar-rmel* existente en la otra orilla del Estrecho para, posteriormente, conforme se empieza a recoger la información de forma escrita y el romandalusí sea una realidad (pensamos que rescatando en parte su denominación romana), pase a ser denominado simplemente *Wadi lakko* (*lacus*), formando esta voz híbrida al unirse los términos río, del árabe, y lago, del iberorromance. Aunque en un nivel más culto fuese conocido como *Wadi ar-rmel lacus quoque piscatorios et portus*, o de manera similar, que derivaría, así lo propone-

mos como hipótesis, en *Wadi ar-rml lac*[*q*]*a* para, finalmente, quedar de forma contraída y por efecto del fenómeno de imela en *Wad ar-rmlaqe* > Guadarranque. Sostenemos que así era conocido este río durante el siglo VIII y buena parte del IX.

Hay otro posible origen más simple, aunque menos probable, y quizá por ello no tenido en cuenta hasta ahora, pero que creemos interesante señalar; es el de *Wadi Arranque*, río de la desbandada, si nos atenemos a una acepción antigua del término «arranque», o «río del comienzo», en el sentido que tiene «arranque» en la terminología militar. Pero, de cualquier forma, sería un hidrónimo surgido muy posteriormente a los sucesos de 711. Lo podríamos datar en la Baja Edad Media, cuando por «arranque» se entendía el lugar desde donde se empezaban a disponer las columnas de los ejércitos y también lugar de «vencimiento de tropas enemigas»[45], quizá en recuerdo del trascendental enfrentamiento entre Tariq y Rodrigo.

A MODO DE CONCLUSIÓN

A pesar de lo señalado en estas páginas sobre la importancia y la necesidad que tiene el estudio de la toponimia arabo-bereber en los inicios de al-Andalus, debemos advertir que su utilidad como fuente de información en la investigación histórica no ha sido suficientemente tenida en cuenta.

Filólogos y lingüistas tienen por delante una ardua tarea en este sentido, puesto que la toponimia de la zona del cono sur peninsular en época de la invasión apenas se ha estudiado y es un recurso imprescindible para la investigación. En este caso es crucial el trabajo interdisciplinar para avanzar y extraer conclusiones fiables, aunque bien sabido es que el estudio toponímico presenta serias limitaciones y condicionantes, de tal modo que no debe utilizarse como única fuente.

Para finalizar, y a la espera de las conclusiones de varias investigaciones, en las que se incluye una revisión de las traducciones de las fuentes árabes, no nos atreveríamos a

dar por zanjada esta cuestión, pues todavía queda pendiente el estudio de un gran número de topónimos, sobre todo menores, que parecen estar relacionados con los sucesos de principios del siglo VIII. Su estudio ofrecerá sin duda una importante información que consideramos fundamental para el conocimiento de los inicios de al-Andalus.

BIBLIOGRAFÍA CITADA

-*Ajbar Maymu'a,* (Colección de tradiciones), Emilio LAFUENTE (Trad.). Madrid, Guillermo Blázquez, 1984.

-AL-MAQQARI *apud* SEGURA GONZÁLEZ, Wenceslao, «Inicio de la invasión árabe de España. Fuentes documentales», *al-Qantir*, (Tarifa), nº10, (2010).

-BARBERO, Abilio, «El Pensamiento político visigodo y las primeras unciones regias en la Europa Medieval», *Hispania*, nº115 (Madrid, 1970), pág. 245-326.

-BENEROSO, José, «Acerca de la entrada de los arabo-bereberes en la Península Ibérica en el año 711. Hipótesis, ucronía, y realidad histórica.», *Almoraima*, 36 (abril 2008), pág. 129-138.

-ÍD., *La incursión de Tarif ibn Malik en 710. Preludio de una invasión* (Nueva Edición completa y revisada). Saarbrücken, Editorial Académica Española, 2012.

-ÍD., «Debate historiográfico e interpretativo en cuanto al enfrentamiento entre Tariq y Rodrigo en julio de 711 (I)», *Almoraima*, 53 (2020), págs.19-26.

-BOUALLAL, Karima, «El latín y su influencia en el bereber», *Onomázein*, 27 (junio 2013).

-CORRIENTE, Federico, *Diccionario de Arabismos y voces afines en Iberorromance*. Madrid, Editorial Gredos S.A., 1999.

-CHALMETA, Pedro, *Invasión e Islamización*. En Colección al-Andalus. Madrid, Editorial Mapfre, 1994.

-DE ALDRETE, Bernardo, «Vocablos godos», en Gregorio MAYANS Y SISCAR, *Orígenes de la lengua española*. Madrid, Librería de Victoriano Suárez, 1873.

Disponible en (6-7-2021):

https://www.bvfe.es/es/directorio-bibliografico-diccionarios-vocabularios-glosarios-tratados-y-obras-lexicografia/16333-vocablos-godos-que-tenemos-en-el-romance.html

-*Digesto* 50
Disponible en (15-02-20)
https://droitromain.univ-grenoble-alpes.fr/Corpus/d-50.htm

-GARCÍA VARGAS, Enrique *et alii*, «Perspectivas de investigación sobre puertos y fondeaderos en el sur de Hispania», en Anna GALLINA y Rita TURCHETTI, *Le strutture dei porti e degli approdi antichi. II Seminario Roma-Ostia Antica 16-17 abril 2004*. Soveria Mannelli, Rubbettino Editore, 2004.
Disponible en (18-02-20):
https://books.google.es/books?id=UPn5puDa_hwC&pg=PA10&lpg=PA10&dq=Carteia+lacus+quoque+piscatorios+et+portus+in+censum+dominus+debet+deferre&source=bl&ots=2ehKiIcU0J&sig=s22Mqr3iDdtWtNGesy-TecyUPM&hl=es&sa=X&ved=0ahUKEwiLu__7o7PJAhVFwBQKHfdbAz0Q6AEIHTAA#v=onepage&q=Carteia%20lacus%20quoque%20piscatorios%20et%20portus%20in%20censum%20dominus%20debet%20deferre&f=false

-GUICHARD, Pierre, *De la Expansión Árabe a la Reconquista: Esplendor y Fragilidad de al-Andalus*. Granada, El Legado Andalusí, 2002.

-HERNÁNDEZ DEL PORTILLO, A., *Historia de Gibraltar*. Ed. de A. TORREMOCHA. Algeciras, UNED, 1994.

-LÉVI PROVENÇAL, E., *La Peninsule Iberique au Moyen Age*. Leiden, E.J. Brill, 1938.

-MARCOS-MARÍN, Francisco, «Notas sobre los bereberes, el afrorrománico y el romance andalusí», *Hesperia, Culturas del Mediterráneo* (junio 2015) pág. 203-221.

-ÍD., «Latín, beréber, afrorrománico, iberorrománico y romance andalusí. Interacción, desaparición y pervivencia de lenguas», *Revista Iberoamericana de Lingüística*, 10 (2015), págs.33-91.

-RECIO ESPEJO, José Manuel, «Medio natural y cuaternario de Gibraltar en los textos de viajeros por la península Ibérica durante los siglos XVIII y XIX», *Almoraima*, 35 (2007), págs.11-20.

-SANZ SERRANO, Rosa, «Aproximación al estudio de los ejércitos

privados en Hispania durante la antigüedad tardía», *Gerión*, 4 (1986), págs. 225-264.

-TORO LILLO, Elena, «La invasión árabe. Los árabes y el elemento árabe en español».

Disponible en (6-7-21):

http://www.cervantesvirtual.com/obra-visor/la-invasin-rabe-los-rabes-y-el-elemento-rabe-en-espaol-0/html/00b64db8-82b2-11df-acc7-002185ce6064_2.html

-VALLVÉ, J., «Sobre algunos problemas de la invasión musulmana», *Anuario de Estudios Medievales*, nº 4 (1967), págs. 361-368.

-ÍD., *La división territorial de la España musulmana*. Madrid, CSIC, 1986.

-VILLAVERDE VEGA, Noé, *Tingitana en la antigüedad tardía (siglos III-VII): auctonía y romanidad en el extremo occidente Mediterráneo*. Madrid, Real Academia de la Historia, 2001.

-VOLNEY C.F., *Viage por Egipto y Siria durante los años 1783, 1784 y 1785*. Trad. Nicolas LLUY, t. II. Paris, Imprenta de Julio Didot, 1830.

Disponible en (18-02-20).

https://books.google.es/books?id=518LAAAAYAAJ&pg=PA210&lpg=PA210&dq=Viaje+por+Egipto+y+Siria&source=bl&ots=WAbreb1k8u&sig=g_RDr4llQAwMB8rM6WLf9Al4m1M&hl=es&sa=X&ved=0CEAQ6AEwBGoVChMIn56ftLD1yAIVQn0aCh1sYQgm#v=onepage&q=Viaje%20por%20Egipto%20y%20Siria&f=false

NOTAS

1 Francisco MARCOS-MARÍN, «Latín, beréber, afrorrománico, iberorrománico y romance andalusí. Interacción, desaparición y pervivencia de lenguas», Revista Iberoamericana de Lingüística, 10 (2015), págs.33-91
2 Este cambio es un proceso evolutivo que culminará ya en el siglo X.
3 Noé, VILLAVERDE VEGA, Tingitana en la antigüedad tardía (siglos III-VII): auctonía y romanidad en el extremo occidente Mediterráneo. Madrid, Real Academia de la Historia, 2001, pág. 307.
4 Idem
5 Estos pueblos, visigodos y vándalos, conservaron una relación más estrecha de lo que habitualmente se ha venido considerando.
6 Francisco MARCOS-MARÍN, «Notas sobre los bereberes, el afrorrománico y el romance anadalusí», Hesperia, Culturas del Mediterráneo (junio 2015) pág. 214.
7 Francisco MARCOS-MARÍN, «Latín, beréber, afrorrománico, iberorrománico y romance andalusí: Interacción, desaparición y pervivencia de lenguas»…, pág.54.
8 CHALMETA, Pedro, Invasión e Islamización. En Colección al-Andalus. Madrid, Editorial Mapfre, 1994, págs. 119-168.
9 Karima BOUALLAL, «El latín y su influencia en el bereber», Onomázein, 27 (junio 2013), pág. 104.
10 Pedro CHALMETA, Op. cit., pág. 126.
11 Pierre GUICHARD, De la Expansión Árabe a la Reconquista: Esplendor y Fragilidad de al-Andalus. Granada, El Legado Andalusí, 2002, pág. 28.
12 VOLNEY C.F., Viage por Egipto y Siria durante los años 1783, 1784 y 1785. Trad. Nicolas LLUY, t. II. Paris, Imprenta de Julio Didot, 1830, pág. 102.

Disponible en.
https://books.google.es/books?id=518LAAAAYAAJ&pg=PA210&lpg=PA210&dq=Viaje+por+Egipto+y+Siria&source=bl&ots=WAbreb1k8u&sig=g_RDr4llQAwMB8rM6WLf9Al4m1M&hl=es&sa=X&ved=0CEAQ6AEwBGoVChMIn56ftLD1yAIVQn0aCh1sYQgm#v=onepage&q=Viaje%20por%20Egipto%20y%20Siria&f=false

13 Federico CORRIENTE, Diccionario de Arabismos y voces afines en Iberorromance. Madrid, Editorial Gredos S.A., 1999, pág. 22.
14 Ibidem, pág. 23.
15 También no sería desacertado hablar de lenguas líbica y bereber.
16 Federico CORRIENTE, Op.cit. pág.18.

17 Es característico el caso del sufijo latino /-ana/, tan abundante en toda la Península, que por imela «inflexionó» a menudo en /-ena/ y a veces en /-ina/.

18 En Elena TORO LILLO, «La invasión árabe. Los árabes y el elemento árabe en español».

Disponible en:
http://www.cervantesvirtual.com/obra-visor/la-invasin-rabe-los-rabes-y-el-elemento-rabe-en-espaol-0/html/00b64db8-82b2-11df-acc7-002185ce6064_2.html

19 Idem.

20 Federico CORRIENTE, Op.cit., pág. 23.

21 Así señala insistentemente el profesor Corriente que debe ser, y no al-Andalus.

22 Federico CORRIENTE, Op.cit., pág. 215.

23 J. VALLVÉ, «Sobre algunos problemas de la invasión musulmana», Anuario de Estudios Medievales, nº 4 (1967), pág. 362.

24 Idem.

25 Vid. Joaquín VALLVÉ, La división territorial de la España musulmana. Madrid, CSIC, 1986, págs. 31 y ss.

26 Vid. Rosa SANZ SERRANO, «Aproximación al estudio de los ejércitos privados en Hispania durante la antigüedad tardía», Gerión, 4 (1986), págs. 225-264.

27 Para más detalles remitimos a Abilio BARBERO, «El Pensamiento político visigodo y las primeras unciones regias en la Europa Medieval», Hispania, nº115 (Madrid, 1970), pág. 288.

28 La presencia de estas sortes ha perdurado en muchos topónimos peninsulares como Suertes, Valdegodos o Villadegodos.

29 J. BENEROSO, La incursión de Tarif ibn Malik en 710. Preludio de una invasión (Nueva Edición completa y revisada). Saarbrücken, Editorial Académica Española, 2012, pág. 48.

30 Cfr. Bernardo DE ALDRETE, «Vocablos godos», en Gregorio MAYANS Y SISCAR, Orígenes de la lengua española. Madrid, Librería de Victoriano Suárez, 1873, págs. 182-184.

Disponible en (6-7-2021):
https://www.bvfe.es/es/directorio-bibliografico-diccionarios-vocabularios-glosarios-tratados-y-obras-lexicografia/16333-vocablos-godos-que-tenemos-en-el-romance.html

31 HERNÁNDEZ DEL PORTILLO, A., Historia de Gibraltar. Ed. de A. TORREMOCHA. Algeciras, UNED, 1994.

32 Vid. José Manuel RECIO ESPEJO, «Medio natural y cuaternario de Gi-

braltar en los textos de viajeros por la península Ibérica durante los siglos XVIII y XIX», Almoraima, 35 (2007), págs.11-20.

33 Ajbar Maymu'a, (Colección de tradiciones), Emilio LAFUENTE (Trad.). Madrid, Guillermo Blázquez, 1984, pág.22.

34 Las formas: lago, albufera, albuhera, al-buhayra, al-buhaira y al-bohaira aparecen con bastante frecuencia en las fuentes, tanto árabes como cristianas.

35 AL-MAQQARI apud SEGURA GONZÁLEZ, Wenceslao, «Inicio de la invasión árabe de España. Fuentes documentales», al-Qantir, (Tarifa), nº10, (2010), pág. 112.

36 Ajbar Maymu'a, (Colección de tradiciones), op.cit., pág.21

37 E. LÉVI PROVENÇAL, La Peninsule Iberique au Moyen Age. Leiden, E.J. Brill, 1938, págs. 92-94 y 180.

38 Pedro CHALMETA, Op.cit., pág. 132.

39 Ibidem, pág. 135.

40 Para Chalmeta, Al-Andalus nunca dejó de ser frontera, y como tal siempre se le consideró. Véase Pedro CHALMETA, Op. cit., pág. 35. Es una afirmación recurrente en su obra.

41 J. BENEROSO, «Debate historiográfico e interpretativo en cuanto al enfrentamiento entre Tariq y Rodrigo en julio de 711 (I)», Almoraima, 53 (2020), págs.19-26.

42 Véase sobre esta cuestión J. BENEROSO, «Acerca de la entrada de los arabo-bereberes en la Península Ibérica en el año 711. Hipótesis, ucronía, y realidad histórica.», Almoraima, 36 (abril 2008), pág. 132.

43 Enrique GARCÍA VARGAS et alii, «Perspectivas de investigación sobre puertos y fondeaderos en el sur de Hispania», en Anna GALLINA y Rita TURCHETTI, Le strutture dei porti e degli approdi antichi. II Seminario Roma-Ostia Antica 16-17 abril 2004. Soveria Mannelli, Rubbettino Editore, 2004, pág. 10 y ss.

Disponible en (18-02-20):

https://books.google.es/books?id=UPn5puDa_hwC&pg=PA10&lpg=PA10&dq=Carteia+lacus+quoque+piscatorios+et+portus+in+censum+dominus+debet+deferre&source=bl&ots=2ehKiIcU0J&sig=s22Mqr3iDdtWtNGesy-TecyUPM&hl=es&sa=X&ved=0ahUKEwiLu__7o7PJAhVFwBQKHfdbAz0Q6AEIHTAA#v=onepage&q=Carteia%20lacus%20quoque%20piscatorios%20et%20portus%20in%20censum%20dominus%20debet%20deferre&f=false

44 Digesto 50, 15.4.

45 Según la RAE forma en desuso de «derrota» (sexta acepción) y aparece como ejemplo: «vencimiento de tropas enemigas».

Disponible en (15-06-20): https://dle.rae.es/arrancada

REFLEXIONES SOBRE ESTRATOS Y CONTACTO: AL-ANDALÚS, AFRORROMÁNICO, AMÉRICA, PALESTINA

REFLECTIONS ON STRATA AND CONTACT: ALANDALUS, LATIN-AFRICAN ROMANCE, AMERICA, PALESTINE

Francisco Marcos Marín
University of Texas at San Antonio. Professor Emeritus
Hebrew University, Jerusalén. Profesor Visitante.

Recibido: 28 de septiembre 2021. Aceptado: 1 de diciembre 2021

Resumen: El objeto de la investigación está repartido entre fenómenos observados en al-Andalús, en el Norte de África, con la peculiaridad de que datos de América y sus lenguas pueden ayudar a esclarecer ciertos puntos y, finalmente, en Palestina. Hay por lo tanto un recorrido histórico y una investigación sincrónica. Ello ha dado lugar a reflexiones que relacionan lenguas y situaciones o procesos que afectaron a lenguas que no se suelen presentar de manera conjunta. Quizás por ello presentarlos aquí pueda despertar algún interés.

Palabras clave: Afro-asiático, apache, árabe, arameo, calco, conmutación de códigos, diglosia, hebreo latín, lengua hablada, lenguas indoeuropeas, Palestina, sustrato.

Abstract: The object of this research is divided between phenomena observed in al-Andalús, in North Africa, with the peculiarity that data from America and its languages can help to clarify certain points and, finally, in Palestine. There

is therefore a historical journey and a synchronous investigation. This has given rise to reflections that relate languages and situations or processes that affected languages that are not usually presented together. Perhaps that is why presenting them here may arouse some interest.

Key words: Afro-Asiatic languages, Apache, Arabic, Aramaic, calque, code-switching, diglossia, Hebrew, Indo-European languages, Latin, Palestine, spoken language, substratum

Durante mucho tiempo y especialmente en los últimos veinte años he tenido ocasión de plantearme problemas diversos relacionados con la conformación de las lenguas y los procesos que se producen por contacto lingüístico. En estas reflexiones se ha tenido que revisar también la conformación de los *estratos* lingüísticos, término que se utiliza sin pretensión específica, sólo en el sentido general de capas superpuestas de elementos y estructuras lingüísticas a lo largo de la historia. Puesto que, durante esos años, el objeto de la investigación se ha repartido entre fenómenos observados en al-Andalús, en el Norte de África, en América, especialmente en los Estados Unidos y, finalmente, en Palestina, ha dado lugar a reflexiones que relacionaban lenguas y situaciones o procesos que afectaban a lenguas que no se suelen presentar de manera conjunta. Quizás por ello presentarlos aquí pueda despertar algún interés. Estas investigaciones tienen el atractivo de estar siempre abiertas, a la espera de nuevos descubrimientos o propuestas metodológicas. Por ello puede permitirse a su autor que las presente más como un ensayo que como una tesis. Las referencias aparecerán en notas[1].

El *marco teórico general* es el que abarca las transferencias de elementos lingüísticos de unas lenguas a otras, es decir, los procesos que introducen o alteran elementos léxicos entre idiomas, palabras o frases, en el dominio del *préstamo* y en el del *calco*. Para los fundamentos teóricos y metodológicos se dispone de bibliografía suficientemente precisa, que evita replantearse discusiones conceptuales.

Juan Gómez Capuz[2] ha insistido en la necesidad de tra-

tar la transferencia léxico-semántica entre lenguas como: a) *proceso* lingüístico: un elemento de una lengua pasa a otra y b) *elemento* lingüístico, generalmente léxico. Como se verá, los dos aspectos tienen cabida en el tratamiento que los autores que se analizarán hicieron en sus estudios. También distingue, dentro de su amplia clasificación, dos tipos generales o amplios y diferenciados del *calco*, el estructural y el semántico.

El *calco estructural* implica el proceso de *sustitución morfémica.* No se incorpora un elemento lingüístico (palabra) desde el otro idioma, sino que se *sustituye: fr. gratte-ciel* > *esp. rascacielos.* Lo característico es que haya correspondencia biunívoca, morfema por morfema.

El *calco semántico* "se limita a alterar las funciones o significados de una palabra simple ya existente en la lengua receptora, generalmente, a través de una extensión semántica."

En esta contribución se planteará en primer lugar el problema del calco semántico como proceso lingüístico y como fenómeno histórico, con base en el análisis de una construcción del español, cuya relación con al-Andalús y con la lengua árabe ha dado lugar a un interesante intercambio de opiniones valiosas. Para hacerlo de un modo concreto se recurrirá a un ejemplo de calco abordado desde diferentes puntos de vista, el origen de las construcciones personales de tiempo en español (y también en portugués): *anochecí rico y amanecí pobre.* En el año 1936 Alf Lombard[3] inició el estudio de estas construcciones con esta afirmación:

> En español los verbos *amanecer, anochecer* no solo significan "hacerse de día / noche", sino también, con flexión personal, "experimentar la llegada del día / de la noche", especialmente con referencia al lugar o las circunstancias: "dónde o cómo encontrarse a la llegada del día / la noche".

Se trata del punto de partida de las discusiones sobre la historia y origen de construcciones como: *Amanecí en Madrid*

y anochecí en Toledo. Amanecí con plata y anochecí sin blanca, Amaneció el campo lleno de rocío, en español, o *amanheci sobre os libros, havendo passado a noite a ler,* en portugués.

Lombard analizó abundantes ejemplos con los materiales que estaban a su alcance en esas fechas. La distribución de las acepciones y usos sintácticos concomitantes sería la siguiente, en rápida síntesis: se situaría en primer lugar (nuestro tipo A) el uso impersonal, *amanece, anochece,* documentado en latín y las lenguas románicas y que no se presta a discusión. El término impersonal es cómodo, como señala Lombard en nota, aunque puedan aparecer ejemplos con *el día* o *Dios* como sujeto. En segundo lugar (nuestro tipo B), con mayor margen de opinión y discusión, habría un uso medio, con reflexivo formal o gramatical, con ejemplos del tipo galorrománico: *il lor anuita en la forest,* que pueden darse en español y portugués y se dan desde los textos antiguos: *les anocheció en el bosque* es posible desde los orígenes del español y, con la forma correspondiente, *lhes escureceu na floresta,* del portugués. En tercer lugar (nuestro tipo C), el español y el portugués (con menos certeza el catalán, que puede simplemente calcar el castellano en este uso) ofrecen desde orígenes ejemplos muy abundantes que Lombard toma del *Diccionario de construcción y régimen* de Rufino José Cuervo o del primer tomo del primer e inconcluso *Diccionario histórico* de la Real Academia Española, publicado en 1933. Por ejemplo en el *Libro de los buenos proverbios* (s.XIII), *"Madre, ¿non vedes las yerbas floridas que amanescen verdes y annochescen secas?* O, en la lengua popular moderna, este fragmento de una canción colombiana: *Me llaman el vagamundo / Porque amanezco en la calle: / Amanezca o no amanezca / Eso no le importa a nadie.*

Encontró Lombard también ejemplos del uso personal en rumano con el verbo *a însera: înserară la un sat,* "(él) anocheció en una aldea". También encontró algunos ejemplos menos convincentes para él en macedo-rumano; pero, en conclusión, rechazó la posibilidad de que iberorrománico y dacorrománico compartieran una construcción como áreas periféricas,

sin ejemplos latinos que explicaran un remoto origen común, señalando la posibilidad de que se tratara de un fenómeno propio del área balcánica, conocida por los abundantes préstamos léxicos y estructurales entre las lenguas que la conforman: albanés, búlgaro, griego y rumano, con sus variantes respectivas. El área balcánica requiere una investigación más detenida, pero con los conocimientos actuales puede esperarse que la construcción esté tan generalizada como en rumano. En búlgaro, por ejemplo, la construcción personal aparece en ejemplos como 'anochecí rico, amanecí pobre' замръкнах богат, осъмнах беден o 'anochecí sano, amanecí enfermo' замръкнах здрав, осъмнах болен[4].

En la segunda parte de su artículo, propuso Lombard que la a clave estaría en el árabe, con muchas construcciones personalizadas con ese valor, especialmente en la forma IV de los verbos: أصبح *asbaha* "amaneció", أمسى *amsà* "anocheció": yafᶜaluhu id̲a ʾaṣbaḥa wa yafᶜaluhu id̲a ʾasmà يفعله اذا أصبح ويفعله اذا أسمى. Literalmente "lo hace cuando él amanece y lo hace cuando él anochece", es decir, "lo hace siempre, lo hace a cualquier hora".

El estudio del árabe en al-Andalús ha sufrido cambios profundos especialmente a partir del último tercio del siglo XX. En síntesis se podría decir que se ha pasado de la perspectiva general de la lengua árabe a la perspectiva particular andalusí. Aplicado a nuestro caso concreto ya no basta con decir que existe una forma IV en árabe en la que se encuentran verbos que pudieran coincidir en sus valores semánticos con los usos personales de *amanecer* y *anochecer*. Hay qué explicar qué tipo de construcciones se pueden encontrar en árabe andalusí y en qué medida justificarían ese calco semántico propuesto. Puede anticiparse ya que la respuesta de los arabistas actuales es negativa: no se trataría de calcos semánticos del árabe, precisamente porque la forma IV del verbo no tiene en árabe andalusí la presencia y el uso del árabe clásico.

Las lenguas semíticas, como el árabe, son idiomas de una compleja morfología, que se desarrolla a partir de las conso-

nantes del radical, mayoritariamente tres. En la actualidad, mejor que el modelo de raíz y esquemas (prefijos, sufijos, infijos, alternancias vocálicas) tradicional, se prefiere el modelo derivacional y se habla más de tema y derivación del tema. En el verbo, además de las modificaciones correspondientes a persona, número, género, aspecto (algunos dirían aspecto-tiempo), voz y modo, hay modificaciones que sirven para expresar diversas alteraciones de significado o de función, reflexivo, incoativo, medio-pasivo, cambio de intransitivo a transitivo, entre otras. Es posible señalar la mayoría de modificaciones de sentido y función que corresponden a las formas; pero siempre hay un margen para otras posibilidades. La forma IV se caracteriza por hamza inicial (sobre alif) en el perfectivo, pérdida de la vocal de la primera consonante radical del mismo y vocalismo u-i en el imperfectivo. El hecho de que el imperfectivo, que sólo tiene hamza (sobre alif) en la primera persona del singular, se marque sólo por el vocalismo, que no se suele representar en la escritura silábica del árabe, hace que con frecuencia la forma II tome el lugar de la forma IV.

En cuanto a su significado, es factitiva, frecuentemente causativa, en lo que coincide con la forma II, aunque con matices. Si el verbo era intransitivo en la forma I, la básica, pasa a transitivo y si era transitivo a doble transitivo. Un uso de esta forma es el *denominativo,* por el que se indica[5] "movimiento hacia un lugar, entrada en un período de tiempo o una acción o estado que tiene lugar o se desarrolla durante ese tiempo". En el caso que interesa aquí, expresa "entrar en un determinado período de tiempo". Se trata de una forma del semítico general, puesto que tiene correspondencias en hebreo (con *ha* como característica formal, no hamza) o siriaco-arameo.

En árabe andalusí[6], además de la sustitución de la forma IV por la forma II que pudiera haberse visto favorecida por la presencia de hablantes sudarábigos, como los yemeníes, de cuyo dialecto es un rasgo, la forma IV se funde con la I. La interpretación de Lombard, en este caso, quedaría descar-

tada, porque no se ajusta a la morfología dialectal. Federico Corriente lo ha expresado en varias ocasiones. Los lectores de textos cultos conocían la forma IV, como las otras formas del verbo, ese punto no se discute. Se trata de que las formas "yo amanezco, yo anochezco" son de uso general en español y portugués y no se pudo dar ese uso general en árabe andalusí, porque no se trata de una fórmula de la lengua hablada.

Desde el punto de vista de la lingüística románica, se debe al Profesor Eugenio Coseriu[7], un repaso completo de las posibilidades, con especial atención a las variantes de su lengua materna, el rumano. En la *NRFH,* en 1961, se planteó la cuestión de la personalización de estos verbos en el mundo románico o en sólo parte de él y su posible origen. Su punto de partida se resume en esta frase: "Es, pues, fundamentalmente la afirmada ausencia de la construcción en las demás lenguas románicas lo que orienta a los estudiosos hacia el árabe."

Coseriu resumió a Lombard en cuatro puntos: a) en rumano habría un solo verbo *(a însera)* comparable con los dos verbos hispánicos *amanecer, anochecer;* b) el rumano *a însera* sería formalmente diferente de los verbos occidentales; c) el uso "personal" rumano, por encontrarse también en griego moderno, albanés y búlgaro, no sería latino sino balcánico; d) el uso "personal" no estaría documentado en latín. Oportunamente, señaló la existencia de los tres tipos de construcciones conocidos: "(A) *le amanece a Fulano* (event. *el día);* (B) *amanezco en Madrid* ("en un paraje", "irgendwo"); (C) *amanezco sano* ("en una situación o condición", "irgendwie")".

Utilizó también Coseriu, como síntesis del estado de la cuestión, el *Diccionario crítico etimológico de la lengua castellana* (DCEC) de Juan Corominas, que resume: "Corominas, de acuerdo con Spitzer, rechaza el arabismo semántico de *le amaneció a Fulano en tal lugar* (tipo A), pero afirma que "hay arabismo indudable en la construcción con sujeto personal *Fulano amaneció* (tipos B y C)". En realidad, habría que corregir, *le amaneció a Fulano* corresponde al tipo B y *Fulano amaneció* al C.

Ni en francés ni en provenzal antiguo encontró don Eugenio pruebas del uso de los verbos correspondientes en los tres tipos de construcciones y tampoco le parece que se pueda postular su uso en las lenguas románicas modernas. En realidad, cuando señala usos similares en provenzal y en español moderno, olvida (y ese tipo de "olvidos" son raros en él) que, en los dos casos, se trata de empleos con pronombre *se,* es decir, que no se trata del tercer tipo de la construcción, sino del segundo. Hace la salvedad, respecto al catalán, de que sí puede encontrarse el uso personal (C) en la variante valenciana; pero pudiera deberse a castellanismo. Convendría estudiar esas variantes valencianas, para asegurarse. No se trata tampoco de un asunto fundamental, porque la etimología es sólo uno de los aspectos por considerar en el análisis de las perspectivas de estudio de estos verbos.

Además de la búsqueda y del concienzudo análisis de los posibles ejemplos románicos, la aportación de Coseriu amplía lo conocido de los usos rumanos de esta construcción verbal. A continuación recuerda que la falta de documentación no se puede tratar como prueba en contra e insiste en un posible origen latino de la construcción, del latín hablado, puesto que no se encuentra en el literario. Ahora bien, en su argumentación sigue apoyándose en ejemplos del galo-románico con pronombre, no se trata de construcciones puramente impersonales, como el tipo (C) de español, portugués y rumano. La forma reflexiva, aunque se considere personal lingüísticamente, no corresponde al concepto activo que la construcción tiene en español y portugués o (sin que ello signifique aceptar que éste sea su origen) en árabe.

Parece innecesario alargar la discusión. Desde el punto de vista de la lingüística románica hay modificaciones que se pueden considerar positivas, como las siguientes: a) en rumano y sus variedades hay más verbos comparables con los dos verbos hispánicos *amanecer, anochecer;*

b) el rumano *a însera* y los otros verbos no son formalmente diferentes de los verbos occidentales; c) la explicación

balcánica del uso "personal" rumano no es concluyente; d) el uso "personal" no estaría documentado en latín.

Queda por tanto un terreno dudoso románico: a) Las construcciones de otras lenguas románicas que Coseriu estudia cuidadosamente son construcciones formalmente reflejas. La construcción iberorrománica personal no va acompañada de pronombre reflejo; b) el uso "personal" no está documentado en latín.

En 1948 había publicado Américo Castro *España en su historia*[8]. Como él mismo diría más tarde a Rafael Lapesa (y los dos me confirmaron esa circunstancia), el descubrimiento de esta interpretación se produjo en medio de tales sufrimientos que hubiera preferido que no se produjera nunca. Se trata de una completa revisión de la historia de España, desde la perspectiva de "cristianos, moros y judíos" y su interacción para la construcción, a partir del siglo XIII, de la noción de España. En la segunda página del prólogo se presenta la idea de que lo que se busca es comprender cómo se forma históricamente una vida colectiva y cómo se conforma su expresión lingüísticamente: "habrá que distinguir con cuidado los abstractos conceptos de cultura y civilización, de las vidas colectivas cuya forma expresan. […] Para hacerse inteligible, la historia habrá de dividirse en unidades biografiables[9]." En la página 223 se precisa esta unidad biografiable como prosopografía al decir que "los hispano-godos hubieron de poner su vivir al compás del Islam a fin de luchar contra él y poder subsistir." En la página 228 se concreta: "En el lenguaje y en todo lo restante, la historia hispana es una realidad *sui-generis,* biselada, que no se entiende sino conjugando lo latino-cristiano-europeo con lo islámico-judaico."

Coherentemente con su visión de que "los problemas de lengua escapan a las manipulaciones externas de la gramática llamada histórica, una historia de las cortezas más bien que del meollo de la lengua", se planteó el modo cómo los hablantes de español y de portugués se pueden incluir dentro de una circunstancia externa (*amanecí / amaneció), i*ncluyó la propuesta de Lombard de calco semántico de la forma IV

del árabe en estas construcciones y la situó dentro de su nueva visión de la historia de España:

> Debemos a Alf Lombard el descubrimiento de que tal uso verbal procede del árabe. El verbo *sabaha,* en su cuarta forma *asbaha,* significa 'ponerse de tal o cual manera por la mañana', como si dijéramos, vivir la mañana. [...] Se trata de otro caso de seudomorfosis —o de injerto—, de forma románica y vivencia árabe, [...] el injerto árabe se produce, no sobre nociones objetivas *(hidalgo,* etc.), sino sobre la misma experiencia interna, sobre la forma de comportarse interiormente al expresar la existencia de una realidad objetiva.

Lo que se quiere resaltar en estas páginas no es que Lombard, a quien sigue Castro, tuviera razón, ya se ha visto la crítica arabista. Tampoco la tenía en el caso de *hidalgo,* desde el punto de vista de la gramática histórica. Don Américo tenía un conocimiento básico del árabe y nunca hizo propuestas basadas en él, sino que siempre buscó la ayuda de arabistas profesionales[10]. Recuérdese que, en esa segunda etapa de una vida ya magistral, no le interesaba "una historia de las cortezas", sino que su perspectiva era diferente.

La *NRFH* publicó en 1949 una nota de Leo Spitzer sobre éste y otros fenómenos lingüísticos, seguida, a continuación, en el mismo volumen, del texto de Castro , "Respuesta a Leo Spitzer"[11]. En él precisaba: "Como el enlace entre el sujeto y el objeto de una experiencia es funcional entre hispanos en forma distinta a como lo es entre franceses, la lengua española y la portuguesa van de 'le anocheció' y 'le amaneció' a 'anochecí enfermo y amanecí sano'. No trato de tal fenómeno como de un 'hecho' lingüístico, sino como de un aspecto, entre cien otros, de la funcionalidad característica de una estructura histórico-vital".

La síntesis de la interpretación histórica se recoge en esta oración: "Lo objetivo del conocimiento y la experiencia vivida

ahora por la persona se reciprocan como un dentro y un fuera que se alternan". Es decir, *amanecer* o *anochecer* son hechos objetivos. Encontrarse en ese momento es una circunstancia de la persona. La capacidad para alternar el ir de dentro a fuera y de fuera a dentro le parece una característica de la manera árabe de comportarse ante el mundo. Lo que le importa es esa característica, no el detalle para él importante, pero superficial (la "corteza"), de su origen concreto.

Castro tuvo noticia del artículo de Coseriu, al que respondió en el *Boletín de la Real Academia Española,* en 1966[12]. Lo entendió como un notable ejercicio de análisis puramente lingüístico, como explicó con la claridad que le caracterizaba:

> Añade el profesor Coseriu, con buen sentido, que eso "no afecta en nada la eventual significación [de ese fenómeno] con respecto al modo de vivir y pensar hispánico" que yo le doy, porque esa "significación se encuentra, evidentemente, en otro plano y no en el mero origen de las tradiciones lingüísticas".

Don Américo se situó en este nuevo artículo ya en el marco conceptual expresado desde 1954 por *La realidad histórica de España,* es decir, por la renovación y precisión de las ideas expuestas en 1948 por primera vez. Matizó su pensamiento al explicar que: "El uso personal de *amanecer* y *anochecer* ha ido personalizándose cada vez más, y está ajustado al sistema expresivo de los españoles; esos verbos se usan sin dificultad en todas las personas y todos los tiempos verbales, lo cual nunca ha ocurrido en francés, provenzal o italiano. [...] Así, pues, aun cuando en latín hablado se hubiesen conjugado personalmente los verbos "amanecer" y "anochecer" *(cosa que no sucedió),* eso no habría impedido a castellanos y portugueses dejarse penetrar por usos árabes, vivos y presentes en muchos modos durante siglos."

Quizás no ocurrió así en este ejemplo desde el punto de vista estricto de la gramática histórica; pero el desarrollo de

Firefox about:blank

los estudios andalusíes ha ampliado mucho el número y el cariz de los ejemplos de contacto de lenguas y de interacción entre los estratos lingüísticos a lo largo de muchos siglos. Habrá que quitar algunas cortezas.

Otros resultados y sus tipos pueden apreciarse en el mundo del latín africano, el bereber y los contactos entre las lenguas de la *Africa prouincia* de los romanos[13]. Se puede analizar la relación entre el bereber y el latín a partir del léxico, de los préstamos del latín al bereber[14]. El préstamo léxico es un proceso complejo, en el que se mezclan aspectos materiales con aspectos espirituales. Los primeros derivan de los objetos desconocidos previamente o desarrollados de modo distinto que una lengua toma de otra. Los segundos derivan de los tipos de pensamiento y categorización de la realidad inmaterial que pasan de una lengua a otra. Los préstamos materiales corresponden a palabras como nombres de animales, plantas, frutas, instrumentos, fenómenos físicos o vestidos, como en español *tomate, aguacate,* del nahua, o *palta,* del quechua, o *zaguán* o *azul,* del persa a través del árabe o *acequia,* de esta última lengua. Los espirituales están relacionados con conceptos de la vida intelectual, cultural o religiosa, como *iglesia,* del griego a través del latín, o *causa,* de esta última lengua; pero incluyen también los nombres de oficio, en los que se producen innovaciones derivadas de variaciones técnicas o de consideraciones socio-económicas, así *albéitar* dio paso a *veterinario.*

En la consideración diacrónica del préstamo hay que tener en cuenta un aspecto práctico, del uso en la vida diaria y otro lingüístico. El aspecto práctico se refiere a las palabras que dejan de tener uso porque los objetos ya no tienen empleo y afecta tanto a las palabras del léxico material como del espiritual. En español han dejado de usarse muchos arabismos tanto del ámbito material, por ejemplo *alcoba,* con el sentido de 'peso o balanza pública', distinto del usual de 'habitación, cuarto', o *atijara,* 'precio, salario', o *almaizar, almaizal,* 'tipo de velo o toca', como del espiritual, así los términos técnicos coránicos como *azora* y *aleya;* incluso el término tradicional

116 de 149 27/06/2022, 17:25

para el texto, *Alcorán,* ha sido sustituido por el más general en las lenguas occidentales, *Corán.* En los préstamos del latín al bereber caben las mismas consideraciones. En ambos casos hay que tener en cuenta la posibilidad de un nuevo *préstamo sustitutorio*, procedente de otra lengua con la que se haya tenido contacto posteriormente y que haya implicado un cambio cultural, por ejemplo de religión. Es el caso de la adopción por los bereberes del cristianismo, primero, con su terminología greco-latina, y del islam, después, con sus términos árabes. El préstamo sustitutorio también se aplica a los objetos materiales, que han podido ser sustituidos por versiones más modernas de los mismos y cambiar su denominación tomando un préstamo de la lengua moderna. Por ejemplo, en español el arabismo *alfayate* fue sustituido por otro préstamo, el galicismo *sastre* y, de modo similar, en bereber el arabismo *sayyāra* fue sustituido por el galicismo o hispanismo *ṭomóbil.*

Más interesante, al menos gramaticalmente, cuando se trata de lenguas tipológicamente tan distintas como el bereber y el latín, es la adaptación del préstamo a la estructura de la lengua que lo recibe. Es un fenómeno bien estudiado en otras lenguas, como, en América, el español y el apache chiricahua. De los préstamos del español al apache chiricahua, hay dos que alteran el esquema normal de la fonología de esta última lengua. En ambos casos se trata de adjetivos españoles que pasan a verbos en chiricahua[15]: *lô.gò (< loco)* y žî.gò (< *rico).* El apache chiricahua no ofrece, en sus vocablos propios, casos de fonemas */l/* /ž/ en posición inicial, sólo aparecen en las posiciones mediales y finales, de manera que la presencia de estos fonemas en posición inicial indica que se trata de un préstamo; pero la adaptación gramatical es mucho más compleja.

Como el chiricahua carece de una clase morfológica 'adjetivo', estos dos préstamos han pasado a la categoría que cumple las funciones predicativas, el verbo. Su estructura encaja perfectamente con el esquema habitual del verbo chiricahua, son palabras bisílabas y admiten la inflexión verbal.

Se introducen en la lengua como la forma no marcada, que es la tercera persona del imperfectivo disjunto, en la que las sílabas iniciales, *lô* y žî, se han reinterpretado como prefijos temáticos que exigen el tratamiento del imperfectivo disjunto. En este imperfectivo se introduce un elemento *hi-* entre el prefijo temático y los prefijos paradigmáticos. En ambos casos la segunda sílaba, *gò,* se ha reinterpretado como la "raíz". El ejemplo más claro se aprecia en la primera persona del dual del imperfectivo: *lóhì.gò,* žíhì.go, ('nosotros dos estamos locos', 'nosotros dos somos ricos') respectivamente. Estas adaptaciones estructurales hacen muy difícil el reconocimiento de un préstamo y así ocurre entre latín y bereber.

Tampoco hay que excluir la posibilidad de que un supuesto préstamo latino al bereber proceda de otra lengua que se lo haya prestado a ambos idiomas. Schuchardt cita el caso de un tipo de encina, *aesculus,* que es *ickir* o *tickirt* en bereber y que aparece en vasco con la forma *ezkur,* para señalar que se puede tratar de un préstamo latino en todos los casos; pero también de una palabra procedente de un fondo léxico mucho más antiguo, como podría ser el ibérico o el líbico bereber.

Brugnatelli señaló como una de las causas principales que ocultan la importancia de las aportaciones latinas al bereber el que una parte del léxico que se ha llamado "espiritual" anteriormente, como el del espíritu y la religión, se ve afectada por la coincidencia de intereses de la romanización y la arabización, vinculados a las religiones cristiana y musulmana, respectivamente. Sorprende el escaso número de vocablos de origen latino pertenecientes al léxico del espíritu que recogió Schuchardt. Uno es bastante claro, *taghawsa* (< lat. *causa*) y otro más dudoso, *tanumi* «norma, costumbre", que relaciona, con muchas dudas, con el griego *nómos*. Este casi completo vacío de Schuchardt, inexplicable en una región con una tradición de escritura religiosa tan grande como el Africa romana, ha sido parcialmente corregido por Naït-Zerrad[16]. Estos tres términos se pueden añadir al léxico religioso latino:

a) lat. *merced(em)* > tuar. *émerkid* , *emârked* "recompensa divina", con alteración semántica en Ouargla

(*amerkidu* "limosna"), en el Mzab (*amerc'idu* "gracias, recompensa"), en Ghat (*amarkidu* "recompensa por las buenas obras"), en Marruecos central (*bu-imercidan,* nombre del séptimo mes islámico, que corresponde a *bu-lajur'* "el de las recompensas".

b) lat. *peccatum* > tuar. *abekkâd'* "pecado", cabilo *abekkad'(u)* "mal/enfermedad; gran desgracia".
c) lat. *angelus*, gr. *angelos* > tuar. *äng›elus* , *ängälos*, "ángel", Ghadames *äng›alus* "inspiración, espíritu", Mzab *äng›elus* "niño pequeño e inocente".

Este número escaso puede incrementarse con una más completa investigación; pero también teniendo en cuenta lo que podría llamarse las *huellas léxicas* en la sustitución de vocablos latinos por árabes. En efecto, por debajo del árabe es posible atisbar huellas del latín anterior, gracias a índices morfológicos. Lewicki y Brugnatelli apuntaron algunas de las posibilidades. Por ejemplo, son femeninos tanto el español *mezquita* como el bereber *timezgida* < ár. *Masjid,* "mezquita", que es masculino. Se trata de un préstamo antiguo en ambas lenguas y ambas coinciden en un género distinto del árabe. Ese femenino podría explicarse por la pervivencia del género del lat. *ecclesia*, documentado por Lewicki[17] en topónimos como *Tayli:siya*, ciudad del imperio de los Banu Rustam, y *ᶜAyin Taylis* en el Gebel Nefusa. Brugnatelli[18] se pregunta si podría aplicarse el mismo razonamiento para explicar el femenino dominante en las designaciones de la Fiesta Grande y la Fiesta Pequeña del islam: *lâid tamez'yant* y *lâid tameqʷrant*. *ᶜaid* 'fiesta' es masculino en árabe, mientras que *lâid* es femenino en bereber, como prueba la concordancia con las formas femeninas de los adjetivos. El argumento sería que bajo el arabismo *lâid* "fiesta (religiosa)" estaría la huella de un femenino antiguo *tafaska* < lat. *pascha*, con el valor genérico de "fiesta", que no sólo se conserva en español, sino que, se puede añadir, aparece en una de las jarchas en romance andalusí, la jarcha A12/H5, conservada en una muwashshaha de Yehuda Ha-Leví, hebrea, y en una árabe de Abu Bakr Ibn Baqī.

Pasando ahora a la sustitución de elementos, un caso paradigmático es el de los numerales[19], que son préstamos en buena cantidad de lenguas, especialmente a partir de la decena. El bereber ha tomado sus numerales del árabe; pero, para ciertos casos, como el bingo, emplea algunos numerales italianos. Es el caso de *kin,* 'línea', cuya relación con el italiano *cinquina* no se hubiera podido establecer, si no apareciera en serie con dos numerales más fáciles de interpretar, *ambu* y *tirnu.Tarbaât tameqq*^{w}*er'att la ttelewwit'en di texxamt weh'h'ed-sen; al' ambu, tirnu, kin u-m-a(r)a tesled.* «La mayoría jugaba al bingo en una sala en la que estaban solos. No oías más que *ambu, tirnu, kin.*»[20].

Acerca de la adaptación de un préstamo latino a las estructuras gramaticales puede recogerse un ejemplo de Laoust[21]. En su curso sobre el bereber del centro de Marruecos señalaba la dificultad de reconocer los sustantivos latinos prestados al bereber, con la excepción de los meses del calendario solar. Siguiendo a Brugnatelli, se puede afirmar que ni siquiera en esos casos está todo tan claro y precisamente la adaptación de los nombres de los meses, que se supone sencilla, puede dar una idea de por qué, si hay complejidades mayores, puede resultar imposible determinar el origen latino. Schuchardt analizó la palabra *abruri,* 'granizo', que relaciona con el mes de abril, *abrir,* a partir de un texto de Stumme[22] : *yigh ilkim mars, arkullu tmrasent tumz'z'in, ar soffghnt taydart ... ibrir ar tbruraint, yigh brurint ar th'arracent* «Cuando llega marzo [*mars],* marcea [*tmrasent]* toda la cebada, es decir, hace brotar las espigas ... en abril *[ibrir]* "abrilea" [*tbruraint],* y cuando "abrilea" [*brurint]* se endurece». En el Atlas Medio (información verbal de Mohamed el-Madkouri) la palabra para 'granizo' es *tbruri.* El latín *Mars* y *Aprile,* por tanto, parecen haber servido para el nombre del mes, para un verbo con significado metafórico relacionado y para un sustantivo de un fenómeno meteorológico propio de abril en el Mediterráneo.

Sin embargo, las cosas no terminan de estar claras. El mismo Schuchardt, más adelante, habló de una contamina-

Firefox about:blank

ción (*Vermischung*) posible entre varios elementos originales. Entre las formas relacionadas con *abruri,* se encuentran las que añaden un elemento protético, como *zebruri* y *ah'abruri,* que dan pie a la suposición de que se pudiera haber creado otro nombre de mes desde el supuesto préstamo. Se trataría de *febrero,* (en chleuh *xubr'ayr*, en cabilo *fur'ar'* y asimismo *cebr'ar'i* , un mes en el que se explican bien, quizás mejor, los fenómenos de granizo o de nieve, puesto que, en algunas regiones, *tebruri* significa 'nieve').

En conclusión, puede decirse que el proceso del préstamo del latín al bereber se ve afectado por tres interferencias o alteraciones que hay que tener en cuenta: la *desaparición del objeto,* la *sustitución del préstamo* latino por otro más moderno y la *asimilación del préstamo a la morfología del bereber* con la correspondiente alteración formal.

La primera alteración, por tanto, es la desaparición del objeto que se designaba con un préstamo del latín. En este caso, por la falta de documentación, puede que no haya restos y se trate de una suposición simplemente teórica; pero apoyada por la realidad de estos procesos en el conjunto de las lenguas y por la existencia de categorías o grupos de préstamos que certifican esa existencia. También hay que tener en cuenta que el préstamo latino puede haberse conservado en una de las variantes del bereber, aunque se haya perdido en la mayoría, por lo que el mejor conocimiento de la lengua bereber y sus variantes contribuirá, sin duda, a ir incrementando el conocimiento de estos préstamos y ayudará a comprender mejor las relaciones entre ambos pueblos.

La segunda interferencia, mucho más clara, es la de la sustitución del préstamo latino por un nuevo préstamo. El árabe ha sustituido casi todo el léxico espiritual latino-cristiano, como consecuencia del cambio de la religión cristiana por la musulmana en toda la región. Los restos latinos y la consistencia del préstamo en este rubro muestran la veracidad de esta suposición. Pero no ha sido sólo el árabe el introductor de sustitutos, también las lenguas modernas, especialmente el francés y el español, han podido contribuir a estas susti-

121 de 149 27/06/2022, 17:25

tuciones léxicas, con el inconveniente añadido de que, al ser lenguas románicas, puede ser difícil determinar cuándo el préstamo es del latín o del español o italiano, especialmente.

La tercera interferencia, la fusión del préstamo en la estructura morfológica del bereber, es la más interesante para el lingüista y también la más difícil, aunque tiene la ventaja de que, una vez probada, testimonia sin duda la existencia del préstamo y su completa integración en la lengua receptora. La presencia de los tres tipos de préstamos latinos en bereber prueba el estrecho contacto entre las dos lenguas y apoya el bilingüismo bereber-latino, que coincide con otras manifestaciones lingüísticas del latín en el área durante siglos y, naturalmente, con la pervivencia del bereber hasta hoy. Es un campo de investigación que está lejos de estar cerrado y que, además, requiere completarse con el estudio de los préstamos entre el ibero-romance, el bereber y el árabe andalusí.

El bereber y el latín o el árabe y el español ejemplifican contactos entre lenguas durante varios siglos. Para terminar, se presentará ahora un contacto de siglos que especiales circunstancias políticas han vuelto a poner ante la vista, el del árabe y el hebreo en Palestina y. específicamente, en el estado de Israel..

Se tiende a estudiar las lenguas en contacto, la diglosia y el bilingüismo en un marco dual o reducido. Muchas razones lo explican, especialmente la dificultad que existe de observar varios pares de lenguas en contacto en lugares y épocas diversas. Aunque a continuación se presenten una serie de observaciones sobre el contacto de árabe y hebreo, la pretensión del autor va más allá. Se trata de animar a los estudiosos de lenguas en contacto, en regiones ricas en plurilingüismo, como el Pakistán, donde se presentó esta propuesta por primera vez, para que tengan en cuenta estas notas en sus estudios. Al ampliar ahora y coordinar investigaciones previas en una investigación conjunta se puede aprovechar la experiencia de haber trabajado sobre diversas situaciones: 1) el Mediterráneo occidental entre los siglos V y X; 2) al-An-

dalús; 3) el contacto del español con las lenguas indoamericanas; 4) el contacto entre árabe y hebreo en Palestina.

Una definición simple de *epistemología* diría que es la investigación de lo que distingue una creencia justificada de una opinión. En esta investigación la creencia justificada depende por una parte de la observación histórica y por otra de la observación personal o directa del investigador, combina por ello un cierto trabajo de campo de diversos investigadores[23].

Palestina tiene una larga historia lingüística que no se limita ni a la llegada de los árabes ni a la imposición de su lengua en el s. VII. ni por la creación del Estado de Israel en 1948 y la imposición del hebreo.

Las lenguas indígenas palestinas incluyen el cananeo, el hebreo clásico y el arameo, desde la perspectiva de las lenguas afroasiáticas, pero, con una visión más amplia, deben tenerse en cuenta otras lenguas. Para los sustratos, se considerarán los períodos históricos más tempranos. Para los tiempos actuales el análisis reflejará el contacto con otros idiomas. En el primer grupo, el persa, el griego y el latín contribuyeron a la creación de una capa cultural plurilingüe notable, mientras que, en la era moderna, han tenido su influencia el turco, el inglés, el francés, el alemán y el italiano. Hay que incluir también el español, ya sea por el papel de los judíos sefardíes en el nuevo Estado o por la importancia de España en la administración y gestión de la Custodia Terrae Sanctae hasta finales del s. XX.

Los sustratos lingüísticos ofrecen un amplio campo de reflexión y análisis. Los contactos modernos reflejan esa tradición lingüística del intercambio lingüístico y, además, añaden la influencia de la escuela y los medios de comunicación y la situación cotidiana de las personas que son hablantes nativos de árabe y hacen su vida en un entorno de habla hebrea. El contacto lingüístico y el cambio de código son términos lingüísticos que reflejan los tremendos cambios experimentados por los idiomas y sus hablantes, particularmente desde el final de la Primera Guerra Mundial, cambios que han alterado el patrón de uso y evolución del idioma.

El público en general no es consciente del hecho de que, además del árabe y el hebreo, todavía se hablan otros idiomas antiguos en esa parte de Oriente Medio, que algunos de ellos pueden incluso preceder al hebreo en ese suelo, y que todos juegan algún papel en la evolución y conformación del árabe palestino y su desempeño como lengua hablada a principios del siglo XXI. Si bien los lingüistas pueden tener un mejor conocimiento de la situación lingüística en el área, puede ser útil presentar una visión de ese conocimiento desde el origen y desarrollo de esos estudios hasta el estado del arte actual, aunque limitado por el marco del lenguaje hablado para el período moderno. El arameo predominó como lengua hablada y escrita en Siria, Babilonia y Palestina durante más de mil quinientos años, desde los siglos VII y VI a. JC. hasta los siglos IX y X d. JC. y todavía puede encontrarse como lengua hablada desde el norte de Israel hasta el Irán, si bien las guerras del siglo XXI y la persecución religiosa de los cristianos por los fundamentalistas islámicos han hecho emigrar a centenares de miles de hablantes y causado miles de muertes.

El concepto de Arabia también requiere de algunas precisiones. Durante la mayor parte de la historia, más que a la península arábiga se ha referido al área situada al este de Jordania y sur de Siria, desde donde fue extendiéndose y ocupando el lugar de otras lenguas semíticas. El árabe palestino es un subgrupo del dialecto árabe levantino. Se considera dividido en tres grupos: 1) Palestino urbano (más cercano a sirio y libanés); 2) Palestina rural; 3) Beduino (más cercano a los dialectos de Arabia).

En 1953 Uriel Weinreich[24] defininió dos condiciones que permiten la interinfluencia de lenguas: 1) Tiene que haber grandes grupos de personas que hablen estos idiomas dentro de una sola área geográfica. 2) El grupo minoritario debe ser lo suficientemente grande para que los cambios que se originan en el contacto lingüístico se hagan evidentes en su idioma. Las dos condiciones se cumplen en el área y entre las lenguas objeto de estudio.

Según el Israel Central Bureau of Statistics (2009) la población israelí es de más de 7,3 millones de personas. Casi toda esa población habla hebreo, ya sea como primer idioma o como segundo idioma. El hebreo es hablado como lengua materna por más de 3,8 millones de judíos nacidos en Israel, que constituyen casi el 52,7% de la población. Casi 2 millones de personas en Israel hablan hebreo como segundo idioma. El árabe es hablado como lengua materna por los árabes israelíes, es decir, los árabes con ciudadanía israelí. Aproximadamente 1,5 millones de personas en la población israelí son árabes, lo que constituye casi el 20% de la población total. La mayoría de los ciudadanos árabes son musulmanes, pero también hay muchos cristianos y drusos. Los árabes israelíes basan sus puntos de vista sobre el hebreo israelí en las ventajas que tiene para ellos, independientemente de las disparidades sociales, políticas y culturales que separan a los judíos de los árabes y de las actitudes de los árabes hacia los judíos israelíes y viceversa. El hebreo israelí se percibe en Israel como más importante y más prestigioso que el árabe. Hay miembros árabes en el parlamento israelí (Knesset). Son los representantes formales de los intereses políticos, sociales y culturales de los árabes israelíes. Los árabes israelíes suelen trabajar para empleadores judíos o estudian en instituciones israelíes y no al revés; Por lo tanto, el hebreo israelí es el idioma de comunicación de los árabes israelíes con sus empleadores, autoridades locales o profesores.

En estas circunstancias hay que partir de las influencias de estratos lingüísticos en la historia pasada de las lenguas semitas e irse acercando a la situación actual. Históricamente el arameo, lengua semita del norte, más cercana por ello al hebreo, influye en la morfología del árabe palestino en una serie de aspectos bien conocidos:

a) El desplazamiento de los sufijos pronominales -kum> -kon; -hum> -hon: أخوكن (vuestro hermano) أضربهن (los golpeó).
b) La forma de nisba -ānī حقّاني: (honesto).
c) La forma siríaca de disminución: -ōn / -ūn; -ōs / -ūs:سنّون

(diente pequeño).

d) La forma šaf ʕal: شلهب (enrojecer).
e) *la* como marcador de objeto:؟سمِعتَ لأخوك (¿Escuchaste a tu hermano?).
f) Cambio del vocalismo de la tercera persona masc. El pronombre del objeto ه es [ho / o].

Entre el árabe y el hebreo, en algunos casos, los préstamos pueden mantener su estructura fonológica original, como en el ejemplo:

ʔ `ovd-ot el-baladijje
trabajador-F-PL DEF-ayuntamiento
"Trabajadoras del ayuntamiento".

En otros casos se producen alteraciones morfofonológicas, por ejemplo con la adición del artículo definido árabe a las palabras hebreas israelíes y su asimilación a las *letras solares* con la misma regla que funciona en árabe.

ʔil + toXni:t > ʔittoXni:t
Programa DEF, plan-F-SG
"El programa, el plan".

En otros casos el préstamo del hebreo se somete a las reglas de la morfología árabe. Así se encuentran formas plurales árabes habladas de palabras hebreas israelíes:

ramzor + a: t> ramzora: t
semáforo-M-SG + F-PL "semáforos". (M-PL)
maX \ su: m> maX \ asi: m (y también maX \ suma: t con la forma femenina plural)
barricada-M-SG "barricadas". (M-PL)
banSar> bana:Ser
llanta desinflada-M-SG "llantas desinfladas". (M-PL)

La forma dual árabe de palabras hebreas israelíes se encuentra en ejemplos como.

sadi:n > sadine:n
sábana-M-SG sábana-M-DU
"Sábana (ropa de cama)". "Dos sábanas (ropa de cama)".
(a diferencia de: sadin-a:t "sábanas (ropa de cama)" (M + F-PL))

Un verbo de origen hebreo puede tomar los morfemas verbales del árabe

ana halaX fi-l-moni:t
1ª persona-SG go-Qal-PFV-3-M-SG en-DEF-taxi-F-SG
'Tomé un taxi.'
ka:n jitnaheg mli:X \
ser-3-M-SG comportarse-hitpael-PRE-3-M-SG bien
"Solía comportarse bien".

En ocasiones el préstamos del hebreo se adapta también sintácticamente, como sucede en la adición del artículo. El artículo definido en hebreo israelí se adjunta al segundo componente de compuestos, en lugar de al primero. El árabe hablado sigue la convención hebrea israelí a este respecto:

ma?`areXet X\inu:X > ma?`areXet **el**-X\ inu:X
sistema educativo sistema DEF educación
"Sistema educativo". "El sistema educativo".
(Hebreo israelí: maareXet XinuX "sistema educativo")

La tipología sintáctica puede reflejar también esta aproximación, con una tendencia hacia la tipología sintáctica S V O en lugar de la tradicional V S O:

ha:da - l-magen miS na:fi?
este defensor-M-SG NEG útil-M-SG
(AR) (HE) (AR) (AR)
"Este defensor es inútil".

En lo que corresponde a la Semántica, se confirma la tendencia general de que el préstamo tiene a ser más específico. Es el caso de déSe, "cancha de fútbol" (en hebreo tiene los significados de "césped, cancha de fútbol"). Esta palabra en el préstamo al árabe perdió su significado original en hebreo israelí de "césped". En árabe hablado su significado se limitó a "cancha de fútbol".

Los préstamos entre árabe y hebreo afectan también al campo pragmático, como se aprecia en estos marcadores del discurso:

keilu "como" (más usado)
kaze "tal, hasta tal punto".

El árabe wala (lo juro por Dios) es tomado por el hebreo y retomado por el árabe con otro valor:

Árabe > hebreo> árabe "¡Guau! ¡De verdad! ¡No me digas!".

CONCLUSIONES

Se han presentado en estas páginas ejemplos tomados de la documentación histórica junto con otros procedentes de la observación directa de las lenguas. El campo abarcado, geográfica y tipológicamente, ha sido también amplio. La observación directa permite saber más del proceso de introducción del préstamo; pero tampoco responde a la pregunta fundamental, por qué determinados elementos se toman en préstamo y se producen alteraciones fonológicas, morfológicas, sintacto-semánticas e incluso pragmáticas, mientras que en otros casos estas alteraciones no se producen. Los condicionantes sociales parecen tener más peso que las meras estructuras lingüísticas. Nada de esto es nuevo; pero la capacidad de observación y de revisión de datos de que disponen hoy día los lingüistas es muy superior a cuanto se tuvo en el pasado. Falta todavía una metodología convincente.

En los procesos históricos, las respuestas de la corteza gramatical que están a nuestro alcance pueden poner en duda ciertos componentes de las propuestas de interpretación histórica; pero tampoco alcanzan a invalidarlas, porque dichas propuestas se basan en apreciaciones de carácter general y esas generalizaciones pueden prescindir de algunos apoyos concretos, al estar apoyadas por un conjunto mucho más grande de datos. Habría que recordar aquí la afirmación del profesor Coseriu de que la falta de pruebas lo único que dice es que algo no se puede probar, no que sea falso o erróneo.

NOTAS

1 Prescindiré en las citas de detalles bibliográficos innecesarios hoy día o de referencias fácilmente subsanables con consultas a bases de datos bien conocidas y usadas. Poseo copia electrónica de todos los textos citados, lo que resulta muy útil al gestionar las citas, como se sabe. No me encuentro solo en esta decisión.

2 Este autor se ha dedicado plenamente al estudio de los préstamos y fenómenos relacionados y es autor de publicaciones de referencia necesarias. Para estas páginas se remite a su revisión crítica: "El tratamiento del préstamos lingüístico y el calco en los libros de texto de bachillerato y en las obras divulgativas", *TONOS, Revista electrónica de estudios filológicos,* XVII, 2009.

3 "5. Die Bedeutungsentwicklung zweier ibero-romanischer Verba", *ZrPh,* 1936. El conocimiento del alemán de Américo Castro, Eugenio Coseriu y Federico Corriente era excelente, de manera que hay que aceptar que sus interpretaciones de Lombard fueron correctas.

4 Información proporcionada inicialmente por la profesora Eugenia Vucheva y completada por la profesora Dora Mancheva, a quien se deben los ejemplos.

5 William Wright, *A Grammar of the Arabic Language,* I, 1859, pág. 33.

6 Para todo lo relacionado con el árabe andalusí se siguen los escritos de Federico Corriente, sintetizados en *A Descriptive and Comparative Grammar of Andalusi Arabic,* Brill, 2013. Para la relación entre el árabe andalusí y las lenguas románicas es imprescindible el amplio prólogo del mismo autor a su *Dictionary of Arabic and Allied Loanwords. Spanish, Portuguese, Catalan, Galician and Kindred Dialects,* Brill, 2008.

7 "¿ARABISMOS O ROMANISMOS?, ESP. *anochecí en el bosque,* rum. *am înnoptat in padură".* Coseriu tiene en cuenta la nueva versión del libro de Castro, *La realidad histórica de España,* F.C.E., 1954, p. 230 y recoge también la crítica positiva de Rafael Lapesa. Conoce también la crítica de Spitzer y la respuesta de Castro. Presenta su trabajo como una tesis cuyo punto de partida se encuentra en Lombard y que ha sido aceptada por algunos romanistas e hispanistas, como los que se citan.

8 *España en su historia. Cristianos, moros y judíos,* Losada 1948. Especialmente págs. 218-225.

9 *Biografiables* no se refiere aquí a biografías de los personajes históricos. Hoy se hablaría más de prosopografía que de biografía en este caso. Llama la atención la coincidencia con el planteamiento que se citó anteriormente de Ibn Jaldún. No me consta que don Américo estuviera familiarizado entonces con los *Prolegómenos.* No figura en el índice de autores y obras estudiadas.

10 Así, por ejemplo, en sus últimos años, ya en España, coincidió con una

larga estancia de James T. Monroe, quien preparaba su libro *Islam and the Arabs in Spanish Scholarship, 16th Century to the Present* con una beca de investigación, y recurrió a él habitualmente. También recurrió a otras figuras conocidas del arabismo español, puntualmente.

11 El título del artículo de Spitzer es "'Mesturar' y la semántica hispano-árabe", La contestación de Castro se titula "Respuesta a Leo Spitzer".

12 "Sobre «YO amanezco» y «YO anochezco»", *BRAE* tomo XLVI.

13 Marcos-Marin, Francisco (2019). El léxico latino en bereber en el marco del estudio de los romances africanos y el continuo lingüístico andalusí. En Luis Luque Toro, Rocío Luque (Eds.), *Léxico Español Actual VI* (vol. VI, pp. 143-154). Libreria Editrice Cafoscarina. Venezia.

14 Schuchardt, Hugo (1918): *Die romanischen Lehnwörter im Berberischen,* Alfred Hölder para la kaiserliche Akademie der Wissenschaften, Wien. Wagner, Max Leopold (1936): *Restos de latinidad en el norte de África,* Biblioteca da Universidade, Coimbra (libro muy citado, pero del que el único ejemplar que he localizado en España es el mío, puede que haya alguno más en alguna biblioteca privada. Que pueda aparecer en algún catálogo no significa que esté disponible). Rössler, Otto (1962): «Die lateinischen Reliktwörter im Berberischen und die Frage des Vokalsystems der afrikanischen Latinität», *Beiträge zur Namenforschung,* 13, 258-262. Lancel, Serge (1981): «La fin et la survie de la Latinité en Afrique du Nord. État des questions», *Revue des études latines,* 59, 1981, 269-297. Brugnatelli, Vermondo (1999): «I prestiti latini in berbero: un bilancio», *Afroasiatica Tergestina: papers from the 9th Italian Meeting of Afro-Asiatic (Hamito-Semitic) Linguistics, Trieste, April 23-24, 1998,* Marcello Lamberti & Livia Tonelli, eds., Unipress, Padova, 325-332.

15 Hoijer, Harry (1939): «Chiricahua Loan-Words from Spanish», *Language,* 15, 2, esp. 111-112 y 112-114.

16 Naït-Zerrad, Kamal (1998): *Lexique religieux berbère et néologie: un essai de traduction partielle du Coran*, Centro Studi Camito-Semitici/ Associazione Culturale Berbera in Italia, Milano págs. 78, 83, 100, 154.

17 Lewicki, Tadeusz (1958): «Une langue romane oubliée de l'Afrique du Nord. Observations d'un arabisant», *Rocznik Orientalisticzny*. XVII, 415-480.

18 *Op. cit.*: 330.

19 Marcos Marín, F. (1992): «Los nombres de los números ante el préstamo», *Actas del II Congreso Internacional de Historia de la Lengua Española,* Pabellón de España (dist. Arco Libros), Madrid, 1173-1194.

20 Picard, André (1958): *Textes berbères dans le parler des Irjen (Kabylie - Algérie)*, Alger (2 vols.), I, 14.

21 Laoust, Emile (1939) : *Cours de berbère marocain. Dialecte du Maroc central*, Geuthner, Paris, pág. 6.

22 Stumme, Hans (1899): *Handbuch des Schilhischen von Tazerwalt*, J. C. Hinrichs, Leipzig.

23 Fraenkel, Siegmund (1886): *Die aramäischen Frendwörter im Arabischen,* Brill, Leiden; Bassal, Ibrahim (2010): "Strata of Foreign Languages in Palestinian Spoken Arabic in Israel", *LiCCOSEC (The Arabic Language and the Palestinian Folklore in Daily Life. Lingua-Culture Contextual Studies in Ethnic Conflicts of the World)*, 15, Osaka: 3–18 y (2012): "Hebrew and Aramaic Substrata in Spoken Palestinian Arabic", *Mediterranean Language Review,* 19, 85-104; Dekel, Nurit y Hezi Brosh (2012): "Languages in contact: Preliminary clues of an emergence of an Israeli Arabic variety", *Journal of Arts and Humanities (JAH),* 1/3, 1-17.

24 Weinreich, Uriel (1953) *Languages in Contact: Findings and Problems*. Mouton, La Haya.

NECROLOGÍA

JUAN CARLOS RUIZ SOUZA (1969-2021): EN BUSCA DE ORIENTE EN OCCIDENTE

Francisco Marcos Marín
University of Texas at San Antonio

La muerte temprana de Juan Carlos Ruiz Souza supone uno de esos golpes a la investigación original que, en mi caso, será muy difícil, si no imposible, de reparar. Aunque no habíamos tenido tiempo de desarrollar la amistad que yo al menos hubiera deseado, porque nuestro conocimiento personal se remonta a hace unos tres años, su obra ya había llamado mi atención y me animó a escribirle e iniciar una pequeña serie de desayunos en la cercanía de nuestros domicilios madrileños en donde descubrí a un sabio que era además una persona maravillosa. No soy original en ello, si se repasan las notas de dolor de colegas, alumnos y compañeros de investigación, hay una coincidencia total en la capacidad humana de nuestro llorado amigo, desde sus habilidades culinarias a su gratísima conversación. De su capacidad científica, la abundante bibliografía que deja también confirma su excelencia. Juan Carlos queda en el recuerdo de todos como un magnífico conferenciante, un gran profesor y un insuperable compañero de investigación. La sensación de pérdida es terrible. Como voy haciendo en esta triste sección que se ha incorporado a la *Revista Iberoamericana de Lingüística,* se destacará lo que permanece, lo que preserva en el recuerdo la vida de quienes ya no están entre nosotros, puesto que las notas sobre los expedientes académicos y las publicaciones se encuentran fácilmente en otros lugares, accesibles electrónicamente.

Para situar a nuestro recordado maestro, porque lo fue, pese a su juventud, baste resumir que se doctoró en Historia

del Arte por la Universidad Autónoma de Madrid con una tesis sobre arquitectura medieval en la Corona de Castilla y el Reino de Granada en el siglo XIV y que desde 2006 era profesor en el Departamento de Historia del Arte I (Arte Medieval) de la Universidad Complutense. Coincidíamos plenamente en nuestras aficiones viajeras y en el convencimiento de que para aprender es necesario estudiar en diversos lugares que en su caso incluyen, sin limitarse a ellos, el Institut du Monde Arabe de París, la Hispanic Society of América, el Metropolitan Museum of Art de Nueva York, el Istituto Storico Italiano per il Medioevo de Roma o el Institut Français d'Études Arabes de Damasco. Los dos tuvimos la fortuna de disfrutar, lamentablemente por separado, de esa ciudad maravillosa que era Damasco, cuyo recuerdo permaneció imborrable y fue recurrente en nuestras conversaciones. La mayoría de sus publicaciones y trabajos están disponibles en Dialnet y otros repertorios. Las líneas siguientes tratarán de mostrar por qué para un filólogo hispanista o para un arabista interesado en al-Andalús, son imprescindibles.

La atomización de la investigación, por la exigencia imposible de una bibliografía ya inabarcable para cualquier tema, hace que sea cada vez más rara la relación de interdisciplinaridad. En esta etapa de mi vida, en la que afortunadamente estoy libre de exigencias administrativas y de requisitos quinquenales, sexeniales o binarios, me esfuerzo en lo posible por acercarme a todos aquellos que, desde otras disciplinas, se aproximan a lo que va quedando de mi objetivo de investigación desde hace ya casi sesenta años: qué ocurrió en el Mediterráneo occidental entre los siglos V y X, especialmente, si se quiere, entre los siglos VII y IX. Cualquier persona mínimamente familiarizada con el devenir histórico de Hispania pensará inmediatamente en hispano-romanos, bizantinos, visigodos y musulmanes, en términos humanos. En términos lingüísticos el campo puede abarcar el bereber, el latín y sus variantes afrorrománicas e iberorrománicas y el árabe. Para completar el panorama parece natural pensar en qué ocurre desde el punto de vista arqueológico y artístico en el noroeste

de África y en Hispania, las Islas Baleares y Cerdeña.

Cuando un profesional de mi edad busca el magisterio de otro profesional casi un cuarto de siglo más joven, suele hacerlo llevado de una necesidad concreta. En mi caso se trataba de la relación entre las iglesias de doble ábside en el norte de África y en Hispania. El cristianismo hispano tuvo una relación de origen y desarrollo durante muchos siglos con África. La provincia romana de África, cuya capital era Cartago, era el gran centro económico que permitía la subsistencia del Imperio y un centro científico y cultural sólo superado por Roma y, quizás, alguna otra ciudad de Italia. Las Islas Baleares fueron punto intermedio en el trasiego de libros y cartas entre Cartago (piénsese en San Agustín, por ejemplo) e Hispania. Lo que desde el punto de vista de la escritura estaba claro para mi tesis era preciso comprobarlo también en la arquitectura y el arte en general, porque el mundo de los tejidos y el cuero fue también un área de conversación común.

Esta inquietud compartida me proporcionó una bibliografía cuidadosamente seleccionada, que me ahorra, estoy seguro, muchas horas de trabajo innecesarias. Me abrió los ojos a mucho más. Cuando pensaba en el título de esta nota, para que no fuera algo tan frío como un nombre y unas fechas, se me ocurrió que quizás lo que mejor resumía el planteamiento de Juan Carlos no era esa especificación de "En busca de Oriente en Occidente", sino que hubiera convenido mejor "En busca de Occidente en Occidente", sin olvidar Oriente, naturalmente. No lo hice porque, sencillamente, pensé que a primera vista parecería una errata y no era conveniente. Para ello puedo valerme de muchas citas suyas, porque era un motivo recurrente en su planteamiento teórico; pero, aunque sea larga, me quedaré con lo que se afirma en un artículo que publicó con Alexandra Uscatescu en la revista *Goya*, en 2014. La cita está en las páginas 95 y 96; pero tiene repercusiones en páginas posteriores:

> La reciente celebración del centenario de la conquista musulmana ha facilitado una renovada re-

> flexión sobre el desarrollo del arte tardoantiguo hispano y las raíces de las formas medievales. La ingente literatura especializada generada en las últimas décadas obliga, más aún si cabe, a una serena meditación interdisciplinar sobre el tema. Sin duda, los eventos y congresos surgidos a la sombra de dicho aniversario han mostrado, desde múltiples puntos de vista, la complejidad de la cuestión. En casi todos ellos, la fecha del 711 parece haber sido utilizada como punto de inflexión en el ámbito artístico, a veces de forma inmediata, otras de manera más pausada; en cualquier caso, salvo excepciones, el 711 se continúa considerando una frontera entre lo hispano-visigodo y lo musulmán. En las últimas décadas, también se ha abordado el tema desde una perspectiva que ha incidido en dicha frontera, al explicar que la presencia de cualquier forma arquitectónica o elemento ornamental hispano, susceptible de ser datado entre los siglos VI-IX, puede siempre estudiarse bajo la óptica de la influencia de lejanos edificios, lo cual ha provocado el retraso de su datación. Unas propuestas cronológicas que suelen remitir a mediados del siglo VIII, es decir, que coincidirían con la llegada de la dinastía omeya a la Península. Por ello se hace necesario reflexionar sobre la entidad del material constructivo visigodo conservado y más aún si cabe cuando éste se encuentra reutilizado en construcciones emirales de la Córdoba omeya (siglos VIII-IX).

Parafraseando un dicho español, originalmente referido a la actitud clerical o anticlerical de los españoles, cuando se piensa en la actitud de los arabistas españoles sobre al-Andalús se podría decir "o sopa de moros o moros en la sopa". El resultado, sin embargo, por un camino o por otro, resulta sospechosamente coincidente: para todo lo que pudiera pare-

cer relacionado o similar con algo que se encuentra en Oriente, se acaba proponiendo una fecha posterior. Juan Carlos, en cambio, buscó decididamente un modelo común mediterráneo que permitiera explicar que muchas de las características de los llamados tradicionalmente "arte visigodo" o "arte mozárabe" eran independientes de la instauración de la dinastía Omeya en al-Andalús y, de manera mucho más profunda, cómo esa importancia e influencia que se concede al arte omeya o a determinados monumentos de ese arte, en el caso de Hispania, era injustificada e innecesaria. Hay una tendencia entre los historiadores del arte medieval español a retrasar fechas y rasgos de monumentos hispanos para explicarlos por influencia oriental que un estudio más detenido desmiente. Este estudio tiene que tener también en cuenta las especiales circunstancias que se presentan en esos monumentos orientales. Para ello hace falta una experiencia directa, un conocimiento del terreno y una capacidad de relación de lo estudiado que en el caso de nuestro autor fueron excepcionales. Por ello conviene insistir desde estos tristes momentos en la necesidad de mantener esta línea de investigación. Lo expresó con indudable claridad en las páginas 96-97 del artículo citado:

> Así pues, para acometer una primera aproximación al tema es imprescindible, por un lado, revisar el papel del paisaje monumental hispano y, por otro, analizar la producción artística del Mediterráneo oriental, al ubicarse allí la mayor parte de los paradigmas que teóricamente podrían explicarnos todas esas formas arquitectónicas que no se pueden fechar antes de mediados del siglo VIII, según sostiene la hipótesis del "canal de transmisión omeya". No obstante, veremos cómo ambos pilares sustentan la personalidad del rico horizonte artístico preislámico del Mediterráneo occidental, desde Italia a Túnez, pasando por la España visigoda, en la que se conserva la magní-

Firefox about:blank

> fica iglesia de Santa María de Melque, … el edificio que, de una forma u otra, dentro y fuera de nuestras fronteras, ha focalizado desde sus orígenes este rico debate científico ya desde época de Gómez Moreno (…). …pretendemos demostrar que los lejanos puntos en común no son más que el reflejo cultural de la Romanidad que subyace en todas las orillas del *Mare Nostrum*.

El año 711, desde mi punto de vista, es sólo una frontera bélica: en él se produjo una invasión por gentes del norte de África que hablaban bereber y dialectos latinos, dirigidas por mandos orientales, árabes. Estos invasores, muy probablemente, no tenían intención de permanencia; pero, al encontrarse con la facilidad de la conquista, siguieron avanzando hacia el norte y el este. Las consecuencias lingüísticas y artísticas de esta invasión no fueron inmediatas. La islamización y arabización fueron procesos mucho más lentos, reforzados, sin duda, por la llegada de los Omeyas que huían de Damasco. El mundo inmediatamente posterior al 711 siguió siendo un mundo latino, en el que, como se observa claramente en la numismática, se fueron introduciendo pausada y continuamente elementos árabes e islámicos. En la parte que le concernía, la propuesta de Juan Carlos Ruiz Souza fue que no hubo una llegada innovadora de elementos artísticos orientales, sino que hubo una continuidad porque muchos de esos elementos, que pueden estar documentados también en Oriente, existían ya en Hispania y no hay necesidad de retrasar sus fechas de construcción para hacerlos coincidir con las hipótesis orientalistas, tanto en su versión tradicional (la de la España musulmana) como en su versión reformista (al-Andalús explicado sólo desde Oriente). Sin duda hay que recordar las acertadas palabras de María Jesús Viguera en su trabajo «Al-Andalus: de Omeyas a Almohades», "El estudioso de al-Andalus debe atender cuanto pueda a lo ocurrido en el conjunto árabo-islámico, en cuyos centros clásicos —sobre todo orientales— se crearon los modelos políticos y cultu-

140 de 149 27/06/2022, 17:25

rales a continuación reproducidos en los confines andalusíes, y en general del Garb u Occidente islámico"; pero hay que recordar igualmente que antes del fin del califato omeya y su sustitución por el abbasí en 750, con el traslado de la capital de Damasco a Bagdad y otros cambios para nada banales, Hispania estaba plenamente integrada en el entorno artístico y cultural del mundo mediterráneo. Por ejemplo, como los trabajos del profesor Ruiz Souza han dejado claro, uno de los principales monumentos del arte omeya: Quṣayr ʿAmra, en Jordania, excavado por arqueólogos españoles, no es la fuente de elementos del arte omeya andalusí, sino que comparte con ellos y con otros edificios artísticos del norte de África, una serie de componentes cuyo origen y desarrollo hay que buscar en ese estrato cultural común que fue la cultura del imperio romano. Por ello, retrasar fechas de monumentos hispanos, cristianos o andalusíes, para que se ajusten a las de Quṣayr ʿAmra es innecesario.

Imposible hacer referencia a todas las aportaciones, innovaciones y precisiones de tan llorado amigo en los múltiples campos que tocó su pluma. Quisiera cerrar este recuerdo con una mención a otro tema de interés común, el del arte y producción textil, que también situó en ese espacio mediterráneo en el que tan cómodamente se movió. Así lo muestra, entre otras aportaciones, en su artículo "De las lorigas de cuero a la Tienda del Encuentro. Arquitecturas de propaganda y victoria en el particularismo medieval hispano". Relacionar convincentemente las lorigas de cuero de lengüetas o escamas con la Tienda del Encuentro de Moisés o con los cimborrios de Zamora, Salamanca y Plasencia es tarea que exige muchos conocimientos poco comunes. Un trabajo de madurez, homenaje a su maestro Isidro Bango Torviso, explícito e implícito, porque la escritura demuestra la profundidad del magisterio.

La transición del cuero a la piedra se observa en buen número de edificios, permite conservar de modo indeleble los valores simbólicos que el cuero presenta de modo más frágil. Tampoco hay que confundir la representación de escamas

con la loriga, porque las escamas forman parte de muchos otros elementos naturales. Esta investigación se mueve cuidosamente por ello entre buen número de posibles trampas que sortea demostrando la madurez de su capacidad de selección, elección y síntesis. La lorigas son iconos de trofeo tras la victoria que aparecen representadas como tales en buen número de obras artísticas arquitectónicas. En la presentación de los trofeos éstos se ordenaban en torno a una loriga que colgaba de un asta, precisamente de ello, de la colocación *sub asta,* 'bajo el asta', viene la palabra *subasta.* El botín del triunfo era vendido posteriormente. La loriga de lengüetas pasó de la iconografía a la arquitectura. Constituyó un elemento arquitectónico. Las escamas de la loriga aparecen como tal elemento en la torre de Juan II del Alcázar de Segovia, en el Castillo de Buen Amor de Villanueva de Cañedo (Salamanca), en la cabecera de la catedral de Palencia, en la fachada de la Colegiata de Aranda de Duero, ya en el siglo XVI, en el remate superior de la fachada de San Pablo de Valladolid. En el mensaje cristiano, esta decoración representa el triunfo de la Vida sobre la muerte. Son muchos los ejemplos citados y documentados. En Salamanca y Plasencia las lengüetas o escamas se usan en los cimborrios en cubiertas compartimentadas y separadas en sectores planos, mientras que en Zamora los sectores son convexos. Hay que suponer que estas tres catedrales, a las que podría sumarse Évora, en Portugal, aunque ésta no se articula en sectores, presentan en sus cimborrios el mismo significado simbólico. Juan Carlos estableció la relación con la Tienda del Encuentro de Dios y Moisés, que se documenta detalladamente en el libro del *Éxodo* y también en *Levítico.* En el texto latino de la *Vulgata* aparece la palabra latina *tentorium,* clara referencia a este lugar, donde se situó el arca de la alianza, constituyéndose en el primer Tabernáculo. Los detalles de la estructura de las tiendas, con referencias a las representadas en el *Libro de los juegos* de Alfonso el Sabio, profundizan en esta descripción. Su valor simbólico es también claro, el triunfo del religión verdadera, la que alberga la Ley, sobre el islam. De

ahí también la presencia del yamur o bola superior de muchos alminares, bajo un símbolo cristiano predominante. Es uno de esos trabajos que, cuando uno termina de leerlos, dan ganas de aplaudir.

Non omnis moriar o, si preferimos Jorge Manrique a Horacio, *nos dejó harto consuelo su memoria.* Seguiremos recurriendo a sus publicaciones, le escucharemos con nuestros ojos y aprenderemos de lo que escribió mientras vivamos. Será ceniza, mas tendrá sentido. Gracias, Juan Carlos, por todo, Dios te bendiga.

ÍNDICE

Artículos

Valverde Olmedo, Beatriz: 5
El orden VS en español: explicación paramétrica para sinohablantes.
The Order VS in Spanish: Parametric Explanation for Speakers of Chinese.

Moreno-McMahon, Sara Cecilia y Yasmine Beale-Rivaya: 27
Trampas en la traducción automática: Un estudio basado en textos periodísticos en línea.
Pitfalls in Machine Translation: A Study on web-based news articles.

Jeon, Yun Sil, Alejandro Muñoz-Garcés, Lindsey Bell and William Jones: 59
Secuencia gramatical para la enseñanza del español como lengua extranjera: nuevas ideas.
Grammatical Sequencing in Spanish as a Second Language: New ideas.

Beneroso Santos, José: 77
Aproximación a la conquista arabo-musulmana de la península ibérica en 711 a través de la toponimia.
Approach to the Arab-Muslim conquest of the Iberian Peninsula in 711 through toponymy.

Marcos Marín, Francisco: 105
Reflexiones sobre estratos y contacto: al-Andalús, Afrorrománico, América, Palestina.
Reflections on strata and contact: Alandalus, Latin-African Romance, America, Palestine.

Necrología

Marcos Marín, Francisco: 135
Juan Carlos Ruiz Souza (1969-2021): En busca de Oriente en Occidente.

Revista Iberoamericana de Lingüística (R.I.L.) tiene como principal objetivo ofrecer las contribuciones y los diferentes estudios científicos en torno a los diversos aspectos de la lingüística en el mundo iberoamericano o desde la perspectiva de los autores que trabajan en ese entorno científico. R.I.L. se publica anualmente. Toda la correspondencia debe dirigirse a la dirección postal *Universitas Castellae* (R.I.L.), Apartado 37, 47080 Valladolid, España. En todo caso es preferible el envío por correo electrónico al editor de la publicación en España (Ricardo de la Fuente) ricardodelaf@hotmail.com. En el correo electrónico, así como en el archivo que se adjunte, deben constar todos los datos del autor: afiliación universitaria y dirección postal y electrónica. Los trabajos que se envíen no deben superar los 30 ff. a doble espacio (incluidas notas y bibliografía), o 7000 palabras. El sistema de referencias debe guiarse por las normas del *MLA Handbook* o seguirá las siguientes instrucciones:

- Libros: Apellido (s), Nombre, *Título* (cursiva) (Lugar: editorial, año) páginas (abreviatura: p., pp.)

- Artículo en revista: Apellido (s), Nombre, "Título" (entrecomillado), *Nombre* de la revista (cursiva), volumen/número (año) páginas (abreviatura: p., pp.)

- Artículo en obra colectiva: Apellido (s), Nombre, "Título" (entrecomillado), Editor (es), *Título* (cursiva) (Lugar: editorial, año) páginas (abreviatura: p., pp.)

- en las notas a pie de página el orden de cualquier referencia debe ser Nombre, Apellido, etc.

No se acepta el uso de subrayados o negritas, sólo las cursivas, salvo que formen parte de algún sistema de notación objeto del trabajo. Tampoco (misma excepción) los signos: < >, «» , { }.

Abreviaturas más frecuentes: *cfr., o.c.* (obra citada), *idem, ibidem, v.* (ver).

Se debe colocar un espacio después de cada signo de puntuación, también antes de guiones (para homogeneizar se recomienda el uso del guion corto).

Para las elipsis debe utilizarse el corchete y tres puntos dentro del mismo […].

Las reseñas darán toda la información del libro, además del número de páginas, el traductor si lo hay, etc. el nombre del reseñador se debe incluir al final del texto.

Para cualquier duda en relación con la publicación, abreviaturas aceptadas, etc. consúltese anticipadamente con el editor de la revista. Los autores interesados en separatas las pueden solicitar anticipadamente (25) y deberán abonar 100 euros.

Los artículos enviados, que deberán ser originales, serán revisados por dos lectores externos a las entidades editoras, que realizarán las oportunas observaciones; en caso de disparidad se enviará el trabajo a un tercer lector externo, y sólo se aceptará el trabajo cuando dos de los informes sean positivos.

Las lenguas en las que los trabajos se pueden presentar son: español, portugués e inglés.

Editado el número, se hará llegar a los autores un ejemplar de la revista.

R.I.L. está incluida en Latindex (directorio y catálogo), DIALNET y está seleccionada para la Base de Datos ISOC.

La suscripción anual es de 20 euros para España y 30 euros para el extranjero. Las personas interesadas en una suscripción pónganse en contacto a través del siguiente correo electrónico: cuc@universitascastellae.es


BASE DE DATOS
ISOC