

Acerca de la transcripción o transliteración del código grafémico árabe al latino, particularmente en su variante castellana

Federico CORRIENTE

BIBLID [0544-408X]. (2002) 51; 361-368

Cuando los usuarios de una lengua, hasta ese momento carente de código grafémico propio¹, o poseedora de uno que deciden abandonar², adoptan el utilizado previamente por otra, o cuando, teniendo código propio, por razones de índole variada tratan de convertir a él otro diferente³, característico de otra lengua o de una tradición grafémica distinta de la misma, se producen determinados problemas de equivalencia entre signos y sonidos, en el primer caso, y entre signos de ambos códigos en el segundo.

Si nos atenemos al último caso, que es el que aquí nos interesa, concretamente restringido a los problemas que suscita la sustitución del código alifático⁴ por la grafía latina, tanto por propósito pedagógico como por conveniencia tipográfica, resulta obvio, en primer lugar que, al tratarse de dos sistemas alfabéticos, o sea, integrados por signos que reflejan aproximada o exactamente el elenco fonemático de las respectivas lenguas, la sustitución puede proponerse tanto a nivel grafémico, lo que produce una transliteración, o sea, una sustitución idealmente unívoca de cada signo del código alifático de la lengua fuente por otro de la grafía latina en la lengua objetivo,

1. Caso, vgr., del latín cuando empieza a escribirse con un alfabeto adaptado del griego, o de éste cuando lo hace con una modificación del fenicio, etc.

2. Por motivos prácticos o ideológicos, o mezcla de ambos, como ocurrió con la sustitución en el rumanol del alfabeto cirílico y en el turco del árabe por el latino, o en el mogol del uigur por el cirílico, etc. Variante especial de esta situación es cuando sólo parte de una comunidad lingüística no sigue el uso mayoritario, sino adopta un alfabeto distinto como señal de identidad peculiar, como es el caso del judeo-árabe, de los aljamiados morisco y judío, del *karšūnī* o árabe escrito en alfabeto siríaco, etc.

3. Caso frecuente por propósito didáctico en una fase inicial del aprendizaje de lenguas que practican códigos muy diferentes, vgr., los del árabe, sánscrito o chino entre los occidentales, aunque también por carencia o carencia tipográfica, o por ir dirigida la información a un público no habituado a dichos códigos, como es frecuente en publicaciones lingüísticas de espectro amplio.

4. Utilizado en principio por el árabe, pero posterior y sucesivamente por otras lenguas del ámbito islámico, vgr., persa, hindí, bereber, turco (hasta la adopción del alfabeto latino), etc., y no exclusivamente por musulmanes, puesto que hay excepciones, como su uso para el hebreo por los caraítas.

como a nivel fonémico, o sea, sustitución de cada fonema representado en la secuencia gráfica alifática por un signo idealmente apropiado para sugerir ese mismo fonema en el elenco de la lengua objetivo. En ambos casos, al no haber correspondencia exacta para ciertos grafemas y fonemas por ellos representados, habrá problemas de equivalencia, a los que habrá que dar algún tipo de solución. Pero la primera cuestión que se plantea, por el hecho de que ni en las lenguas que utilizan el código alifático, ni en las que emplean el latino, existe una equivalencia totalmente unívoca entre grafema y fonema, es, evidentemente, la necesidad de optar por transliteración o transcripción.

La transliteración (o sea, transcripción grafémica) consiste en reflejar cada grafema del texto fuente unívocamente por otro en el sistema grafémico del texto objetivo, de manera que la retroversión pueda ser automática. Es un procedimiento imprescindible en estudios relacionados con cuestiones ortográficas, epigráficas, dialectológicas y, en general, en todos aquellos casos en que, a falta de poder reproducir exactamente la grafía original, conviene que el lector pueda reconstruirla hasta mínimos detalles, pero no es particularmente recomendable en la docencia, donde sólo retrasa la adquisición por el estudiante de las reglas que transforman ciertas secuencias más profundas en otras más superficiales, y es particularmente dañina para lectores no especializados, que suelen ser mayoría en campos como historia y arte, a quienes sugiere pronunciaciones totalmente incorrectas como **Abdalrajmán*, motivada por la grafía ‘Abd-al-Rahmān, bastante más distorsionada que el *Abderramán* de nuestros antepasados.

La transcripción fonémica consiste en reflejar cada fonema del texto fuente unívocamente por un signo en el sistema grafémico del texto objetivo, el cual responderá exactamente a aquél, o bien por existir en ambas lenguas idénticamente, o por recibir una representación especial en la lengua objetivo que permite al conocedor de ambas identificarlo, según procedimientos que serán objeto de posterior comentario. La transcripción fonémica acelera el aprendizaje por los principiantes de las reglas que gobiernan aparentes ambigüedades y equívocos entre grafema y fonema, y permite al profano no cometer más distorsiones fonéticas en la reproducción de las voces alógenas que las inherentes a las diferencias entre su propio sistema fonológico y el desconocido. Volviendo al ejemplo anterior, hiere menos al oído **Abdarrajmán*, y de poco sirve una grafía que teóricamente permitiría reconocer el artículo árabe intermedio, lo que al profano nada dice gramatical ni semánticamente, pero que le hace caer en error fonético⁵.

5. No ignoramos que una de las razones esgrimidas a favor de la transliteración es la claridad con que en catálogos y listados de obras y autores se permite eliminar el artículo árabe de las secuencias alfabéticas,

Recomendamos, pues, la transcripción fonémica, y pasamos a estudiar la siguiente cuestión, a saber, la solución de los casos en que no hay equivalencia exacta para un determinado fonema del texto fuente y, por consiguiente, para el grafema correspondiente en el texto objetivo. Dicha carencia admite tres soluciones:

- a) Invención de un signo nuevo⁶,
- b) Adopción de una convención ortográfica, generalmente basada en agrupar más de un signo preexistente y atribuirle el valor deseado⁷, y
- c) Reutilización de un signo ya existente, generalmente anticuado o poco usado, con el nuevo valor, aunque ello engendre siempre cierta ambigüedad⁸.

Aunque en estricta teoría lingüística, cualquier sistema de transcripción es válido, mientras sea unívoco y claramente definido, la realidad es que la opción entre estas soluciones suele ser gobernada por varios criterios racionales, como proximidad funcional, mayor o menor apego a una tradición grafémica, preocupación por evitar posibles ambigüedades, difusión de determinadas soluciones en áreas culturales más amplias o influyentes o, sencillamente, coste tipográfico.

Volviendo una vez más al caso que nos ocupa particularmente, o sea, la sustitución del código alifático por la grafía latina, es bien sabido que los arabistas europeos del s. XIX se vieron obligados a diseñar sistemas con este objeto, tanto para la divulgación de datos entre un público generalmente desconocedor de las lenguas orientales y sus peculiares grafías, como para la redacción de sus propias investigaciones, en las que el uso de la grafía y posteriormente tipografía no latinas resultaba engorroso, sobre todo por la dirección contraria de la escritura, poco claro por la fácil desaparición o confusión de puntos diacrítico y vocalización extralinear, además de costoso por requerir una complicada tipografía. Hacia mediados de dicho siglo vino a ser bastante común entre los orientistas europeos el uso de representar los fonemas velarizados y la espirante faringal con un punto diacrítico suscrito (**ṭ**, **ḍ**, **ʃ̣**, **ẓ**, **ḥ**), y

según es uso, probablemente práctico y ya irreversible, pero no es menos cierto que las personas que las usan están perfectamente al corriente de ese rasgo, o se les puede informar de él sin mayor problema. Algo parecido se puede decir del generalizado uso de guiones para indicar la segmentación morfológica de la lengua transcrita: de nada sirven al profano, salvo para equivocarle sugiriendo pausas que no existen, ni al especialista, que no necesita le señalen dicha segmentación, por lo que es preferible suprimir tales guiones y sustituirlos por la explicación pertinente en los casos en que su presencia afecta a la catalogación alfabética.

6. Tal fue la invención de **f** y **q** en el alfabeto latino, vgr., aunque es más frecuente la modificación gráfica de un signo ya existente, como es la **ñ** del castellano.

7. Tal hizo el latín al diseñar **th**, **ph** y **th** para reproducir ciertos sonidos de voces griegas, procedimiento que sería muy usado por las lenguas europeas al crear sus ortografías basadas en el alfabeto latino, aunque a veces con equivalencias dispares como se observa en los casos de **ch**, **w**, etc.

8. Caso de la **x** que sustituyó a la antigua combinación **sc** para expresar una consonante chicheante, vgr., en portugués y catalán, sin poder abandonar simultáneamente su equivalencia clásica en otros casos.

utilizar los espíritus suave y áspero de la grafía griega para representar *hamz* y *‘ayn*, mientras que había menos acuerdo en cuanto a la representación de las interdentes, palatales y espirantes velares, de todo lo cual acabarían emergiendo ya a fines del XIX dos sistemas, que básicamente perduran, a saber:

- 1) El centroeuropeo, en el que el alifato árabe se representa con los grafemas **ʾ, b, t, t, ġ, h, h, d, d, r, z, s, š, s, d, t, z, ‘, ġ, f, q, k, l, m, n, h, w, y**, con la adición de las tres vocales (**a, i, u**, y sus homólogas largas **ā, ī, ū** a más del grafema **à**, una concesión a la transliteración que permite la recuperación automática de la *alif maqṣūrah*).
- 2) El anglosajón, en que las interdentes son representadas por **th** y **dh**, las palatales por **(d)j** y **sh**, y las espirantes velares por **kh** y **gh**, digramas a los que, si la tipografía lo permite, se suscribe una raya para evitar la ambigüedad que surge en los correspondientes grupos de la primera letra con **h**.

Cada uno de estos sistemas tiene determinadas ventajas e inconvenientes: el centroeuropeo evita totalmente los digramas y es totalmente coherente, al tener sus diacríticos siempre funciones homólogas (raya suscrita = interdentalización; *háček* o antilambda = palatalización; punto suprascrito = espirantización sonora; punto suscrito = velarización o, en un caso, laringalización; semicírculo suscrito, a menudo sustituido por raya por dificultad tipográfica = espirantización sorda), pero resulta tipográficamente más exigente, mientras que el sistema anglosajón prefiere digramas de equivalencias bien conocidas internacionalmente, si bien su subrayado para evitar ambigüedad viene a anular la economía tipográfica.

La tendencia de ambos sistemas ha sido a consolidarse entre sus usuarios, continentales por una parte, y anglófonos, por la otra, aunque ha habido y sigue habiendo, por supuesto, desviaciones parciales, sobre todo del sistema centroeuropeo, por razones de economía tipográfica⁹, preferencias tradicionales o nacionales e incluso gustos personales, dentro de las cuales merece especial mención, por lo mucho que nos afecta y sus rasgos idiosincráticos, el sistema llamado de la “Escuela de arabistas españoles”.

Para explicar su génesis conviene recordar que, como saben todos los conocedores del tema, el arabismo careció en España de especialistas lingüistas durante su periodo de re inserción tras su liquidación por la Inquisición, que abarca todo el siglo XIX,

9. Entre las modificaciones de esta índole está el uso por algunos orientistas israelíes de los grafemas **j** y **x** por **h** y **ġ**, e incluso **c** por **š**, lo que pareció conveniente suscribir para los dos primeros signos en nuestros estudios de árabe andalusi y romandalusi, no sólo por economía tipográfica, sino por el hecho de que **j** es exactamente el reflejo gráfico posterior más habitual de dicho fonema en los arabismos del iberromance, mientras que **x**, signo adoptado del griego por la International Phonetics Association, supone una solución muy autorizada al problema de transcribir un fonema para el que se ha propuesto demasiadas grafías.

y a lo largo de la mayor parte del de florecimiento, el XX, lo que se suele atribuir al escaso número de investigadores en ejercicio durante todo ese tiempo y a unas muy comprensibles prioridades, tales como la historia medieval, las historias del arte y la ciencia, la literatura andalusí, e incluso la filosofía, el misticismo y la religión. Los cultivadores de estas disciplinas fundamentales se veían a veces forzados a ejercer de filólogos, editando y traduciendo textos, para lo que necesitaban cierto bagaje lingüístico, pero salían del paso de tales retos como podían, con mayor o menor brillantez, sin que hubiera ocasión o deseo de dedicarse más a fondo a tales menesteres, ni proponer a discípulos que lo hicieran, quizás porque ya “lo hacían ellos”, y muy brillantemente, en el resto de Europa. Los que, en palabras de Imru’ Alqays (“¡Cuántas veces del odre de la gente fijé los cabos a mi espalda sumisa y asendereada!”), tuvimos que cargar durante décadas con la responsabilidad abandonada de producir urgentemente inexistentes diccionarios y gramáticas de la lengua árabe para hispanófonos, hemos lamentado siempre ese descuido, del que aún no estamos, ni mucho menos, recuperados. Sin llevar más agua a ese molino, sí diremos que en aquellas circunstancias no es sorprendente que la decisión de Asín en las primeras décadas del s. XX, y ya con considerable retraso con respecto al resto de Europa, de adoptar y, de paso, españolizar la transcripción científica centroeuropea del árabe llevase a resultados bastante extraños y algo menos que recomendables.

Como rasgo positivo único, esta adaptación mantuvo la abstención de recurrir a digramas, pero incurrió, en cambio, en los siguientes graves inconvenientes:

- 1) Anuló el carácter internacional de las equivalencias, al sustituir **h** por **j**, introduciendo una equivalencia propia sólo, en toda Europa, del castellano moderno septentrional, así como crear la peculiarísima **ŷ**, ocurrencia poco plausible.
- 2) No supuso ninguna economía de diacríticos, ya que dicho último signo anula el ahorro producido por la sustitución de **h**, al tiempo que la sustitución de **g** por **g** crea problemas en la transcripción del persa, turco, etc., e incluso de los muchos dialectos árabes, antiguos y modernos, que poseen la oclusiva velar sonora, y hace pensar a los alumnos hispanófonos que tal sea la realización de la fricativa velar sonora, contribuyendo a fomentar de por vida un rasgo adicional evitable de pronunciación incorrecta.
- 3) La pretendida españolización del sistema internacional, auténtica cuadratura del círculo, mediante dichos cambios, se detiene incoherentemente ante **t** y **z**, que hubieran debido de convertirse, según el mismo principio, en **z** y otro signo especial para

la sibilante sonora¹⁰, pero retuvieron aquí por graciosa concesión su valor del sistema imitado, totalmente ajeno al castellano.

Es obvio que la decisión de Asín fue equivocada, debido a su escaso conocimiento de lingüística y no mayor interés por la lengua árabe¹¹, aunque haya sido mantenida hasta la actualidad por casi todos los arabistas españoles, gracias al prestigio de que disfrutaron otros aspectos de su obra y a su dirección monolítica de estos estudios en nuestro país, continuada por sus sucesores y no cuestionada hasta tiempos muy recientes. Obedeció a un prurito de nacionalismo, totalmente inoportuno casi siempre, pero más cuando se pretende operar en terreno internacional, careció de coherencia y produjo un hábito y hasta una tradición, que no hay porqué continuar, del mismo modo que la mismísima y muy conservadora Real Academia de la Lengua Española ha renunciado final y recientemente a la teórica consideración de signo único de **ch** y **ll** en el alfabeto tras siglos de vigencia, contra todos los demás usuarios del alfabeto latino. Además, y por otra parte, tampoco faltaba la conciencia entre los mismos propugnadores de la versión españolizada de la transcripción internacional del árabe de que era de consumo meramente doméstico¹².

Creemos, pues, que no hay ninguna razón, científica por supuesto, ni tradicional a causa de las pocas décadas transcurridas en esa senda errónea, para no abandonar dicho sistema, a medias castellanizado, y sí muchas para regresar al centro europeo en las tareas de carácter científico, poniéndonos al día también en aspectos lingüísti-

10. Hubo de hecho quien, completando coherentemente el proceso, así lo hizo, como E. Ibáñez en sus diccionarios de bereber rifeño y *bā' amrānī*, donde se recurre a **z** y **ç** para ambos fonemas.

11. Hemos demostrado hasta la saciedad estas incontrovertibles afirmaciones, aunque a algunos incondicionales parezcan blasfemia, en nuestro artículo "Las etimologías árabes en la obra de Joan Coromines". En *L'obra de Joan Coromines. Cicle d'estudi i homenatge*. Sabadell, 1999, pp. 67-87, esp. 69-74, no con la voluntad de quitar méritos al maestro, que los tuvo muy notables en otros campos, sino de advertir a incautos del peligro, en que otros ya caímos, de dar fe a sus hipótesis lingüísticas, o de utilizar en la docencia su famosa *Crestomatia*, como aún se hace en demasiados sitios, con efectos a menudo irreparables. Otros disparates que salieron de la misma o parecidas minervas, y que venimos denunciando son las formas **beréber**, **zéjel**, **moax/saja**, etc., por **bereber**, **cejel**, *muwaššah(ah)*, según comentamos por última vez en nuestra obra *Poesía dialectal árabe y romance en Alandalus*. Madrid: Gredos, 1998, pp. 19-20, n. 1. Pero no sería justo omitir que Asín pecó en distinguida compañía, como señalamos en nuestro artículo "Algunas 'palabras fantasma' o mal transmitidas entre los arabismos y voces de origen oriental del DRAE", en *Estudios Árabes dedicados a D. Luis Seco de Lucena*. Granada, 1999, pp. 93-100, por lo que sería deseable, puesto que ninguno está libre de yerro, ni siquiera las academias, que todos estuviésemos un poco más dispuestos a enmendalla, y menos a sostenella.

12. No otra interpretación puede tener el hecho, vgr., de que su discípulo preferido y entusiasta loador García Gómez no la usase, sino se atuviese a la internacional, en obras como su *Todo Ben Quzmān*, para la que buscaba acogida más allá de nuestras fronteras. Lo mismo hizo F. Pareja en la catalogación de la biblioteca, que acabaría llevando su nombre, del llorado Instituto Hispano-árabe de Cultura, la cual, habiendo sobrevivido a éste, es hoy el fondo bibliográfico más importante de nuestro país en este campo.

cos, que ya hace también algunas décadas se cultivan en nuestro país, y aprovechando la coyuntura de las posibilidades que ofrece la informática por lo que a disponibilidad de diacríticos se refiere.

De paso, y ya que estamos hablando de los problemas de la transcripción del árabe, debemos decir algo acerca de la peculiar manera de reflejar la lengua árabe clásica que utilizan muchos arabistas, suprimiendo las inflexiones gramaticales (*i'rāb*) cuando admiten variación, aun en formas contextuales, y manteniéndolas cuando son invariables, aun en formas pausales, con lo que fabrican una lengua que no es árabe clásico, ni dialectal, ni nada que haya nunca existido, sino cómoda corruptela¹³. Concediendo la dificultad de revertir a las formas estrictamente clásicas en los antropónimos y nombres de obras, por las alteraciones que ello conllevaría en catálogos y obras de consulta que han seguido aquella práctica durante decenios, no debería tolerarse tal jerga cuando se hacen citas más o menos extensas de texto árabe clásico.

En otro orden de cosas, sería desaprovechar la ocasión no referirse, aunque sea de pasada, a otros dos aspectos no científicos de la transcripción del árabe, a saber la sistematización de la utilizada en la prensa y obras de divulgación, o de la necesaria para la transmisión de texto árabe por correo electrónico.

En el primer aspecto, es sabido que reina el caos más absoluto, ya que, tras una larga tradición de limitarse a copiar las transcripciones de la prensa extranjera, ajustadas naturalmente a las ortografías de distintas lenguas, aunque finalmente muy dominadas por la norma inglesa de las principales agencias internacionales de noticias, nuestros periodistas han intentado, y lo mismo han hecho los escritores de divulgación, castellanizar topónimos y antropónimos extranjeros de grafía no latina, con resultados tan espantosos como la invención de un “Dien-Bien-Fu”, de los “jemeres” rojos y del imán “Yomeini”. Aun reconociendo que tampoco las lenguas de grafía latina salen mejor libradas de la fonación por nuestras predominantemente monolingües fauces, lo cierto es que no se ha ganado nada, más bien se ha perdido bastante, en esa pretendida castellanización, y que hubiera sido preferible, como se ha hecho en otros países, no adaptar las grafías a la norma local, sino respetar otra de orden internacional, que conocen y aplican las personas cultas, resignándose a que las otras, si no quieren imitarlas, salgan del paso como puedan. Para la grafía árabe que aquí nos ocupa, dicha norma internacional, no científica, insistimos, y por tanto, no

13. Suele verse, vgr., en los títulos de obras, perfectivos en tercera persona del sg. acabados en *-a*, pero nunca un sustantivo en acusativo determinado con idéntica terminación, y es habitual eliminar por convención la *-h* que suple la caída en pausa de *-t* en el morfema femenino, práctica sólo discutible, pero es, en cambio inadmisibles el mantenerla en *Kitāb al-quḍāt bi-Qurtubah* y suprimir cualquier marca en *Risāla Hayy b. Yaqzān*, como hace Asín en las p. 8 y 7 de su *Crestomatia*.

exacta, sino ajustada a la capacidad media de aproximación a la pronunciación correcta de los no especialistas, transcribe el alifato así: Ø (cero para *hamz*), **b, t, th, (d)j, h, kh, d, dh, r, z, s, sh, s, d, t, dh, ', gh, f, q, k, l, m, n, h, w, y**, a lo que se añade la conveniencia de acentuar la vocal tónica, según nuestras normas académicas, para evitar los habituales Mohammed y Yaser (sugerimos Muhámmad y Yásir¹⁴). Se observará el estrecho parentesco de este sistema con el científico anglosajón¹⁵, lo que no es sino una consecuencia del dominio internacional del inglés, que hay que reconocer y aprovechar, dejando toda nostalgia por sistemas más “nacionales”, los cuales, unidos a un necesariamente incompleto conocimiento de todas las lenguas del mundo, desembocan a menudo en disparates como los arriba señalados.

Finalmente, para el envío de mensajes por correo electrónico, en que ningún sistema garantiza hasta ahora la correcta recepción de diacríticos o texto no latino, algunos venimos utilizando un cómodo procedimiento que representa así el alifato árabe: **3, b, t, th, j, H, x, d, dh, r, z, s, sh, S, D, T, Z, 9, gh, f, q, k, l, m, n, h, w, y, a, i, u, aa, ii, uu**, evitando la ambigüedad que pueden causar los digramas mediante un punto separador (vgr., **ra3at.haa, mud.hish, yus.hil**, etc.), aunque también cabe sustituirlos por otras cifras (vgr., **mathal = ma4al, bidhaalika = bi2aalika, mushiir = mu5iir, bighayr = bi8ayr**), o por la adición de una letra no utilizada en lugar de la **h**, vgr., la **I** (*i* mayúscula, o sea, **matlal, bidlaalika, musliir, biglayr**).

Con esta triple propuesta creemos se puede mejorar considerablemente el panorama de la transcripción del árabe al alfabeto latino en sus versiones científica, divulgadora y electrónica, contando con un cierto grado de aceptación, que nunca será total, a causa de preferencias de índole muy diversa.

14. Es cuestionable, aunque preferible, atenerse al sistema trivocálico de la fonología del árabe clásico, evitando el uso de **e** y **o**, siempre problemático.

15. Del que apenas se diferencia sino por la omisión de los rasgos de velarización y faringalización, ajenos a las lenguas europeas, si bien se debe señalar que la mayor parte de sus digramas no pertenecen a la tradición ortográfica inglesa, sino que fueron incorporados tempranamente a un sistema ya internacionalizado y que fue aceptado e integrado en la tradición gráfica de los anglófonos cultos.