

LA INFORMACIÓN LINGÜÍSTICA EN LA TRANSCRIPCIÓN FONÉTICA AUTOMÁTICA DEL ESPAÑOL.

Antonio Ríos Mestre

Laboratorio de Lingüística Informática, Departamento de Filología Española,
Facultad de Filosofía y Letras, Universidad Autónoma de Barcelona.
08193 Bellaterra (Barcelona).

Introducción: algunos problemas que presenta la transcripción fonética automática del español

Un sistema de transcripción fonética automática es un algoritmo que transforma un texto de entrada representado en caracteres grafemáticos en una representación expresada mediante símbolos fonéticos. Dicha transformación puede ser llevada a cabo a través de la aplicación de un conjunto de reglas o a través de la aplicación de un lexicón [1]. La transcripción resultante de este proceso será «fonológica» o «fonética», dependiendo del código utilizado en la salida, y podrá ser interpretada como una representación del texto transformado, en términos articulatorios, acústicos o lingüísticos [2].

Los trabajos sobre la fonetización automática del español [3] han sido concebidos, fundamentalmente, como módulos integrantes de sistemas de síntesis, en los que la transcripción fonética del texto ortográfico procesado es un paso previo para la conversión del mismo en una serie de parámetros acústicos, controlados por el sintetizador, que finalmente serán transformados en voz.

El carácter «fonémico» de la ortografía española [4] hace relativamente fácil esa operación, frente a lo que sucede en otras lenguas, como el francés [5]. Las irregularidades ortográficas: grafemas sin realización fonética (la *h* llamada «muda»), dígrafos (*ch*, *gu*, *qu*, *rr*, *ll*), grafemas con más de una lectura fonética (*c*, *g*, *y*, *i*, *r*) y fonemas representados por más de un grafema (*/i/*: *i*, *y*; */u/*: *u*, *ü*; */b/*: *b*, *v*; */g/*: *g*, *gu*; */k/*: *c*, *qu*, *k*; */ʎ/*: *c*, *z*; */x/*: *g*, *j*; */r/*: *r*, *rr*), son interpretables por el lector español gracias a las restricciones sobre la combinatoria de los mismos previstos por la normativa (p.ej., el dígrafo *qu* sólo aparece ante *e*, *i*, para representar el sonido [k]). También la acentuación fonética de los elementos léxicos es predecible, en principio, a partir de las normas de acentuación gráfica, al tiempo que las restricciones en la combinación de elementos consonánticos y en las agrupaciones de vocales facilitan la organización silábica de las secuencias¹. Los sistemas de fonetización automática del español antes mencionados incorporan en su transcripción la doble información fonológica del acento y la división silábica, así como la realización de variantes alofónicas.

Escapan a la gran regularidad de la ortografía española los extranjerismos no totalmente adaptados a la misma, voces de incorporación relativamente reciente, aún oscilantes en su grafía (*whisky* /*güisqui*) e incluso en su pronunciación, dependiendo del registro de habla (*show*: [ˈtʃow] / [ˈʃow]). En la representación grafemática de este sector del léxico suelen intervenir agrupaciones no existentes en la lengua (*ångstrom*) y grafemas no patrimoniales -como es el caso de la *w*, con una doble lectura, */b/* - */u/*, según la etimología de la palabra [7]- cuya pronunciación, más o menos cercana a la pronunciación de la lengua de origen en función del grado de bilingüismo [8], no es tan fácilmente predecible cuando los términos no están en nuestra competencia. La representación fonética de las irregularidades puede solucionarse ampliando el número de reglas o mediante la introducción de diccionarios de excepciones.

¹ Sobre la pronunciación del español pueden consultarse los manuales citados en el apartado [6] de las referencias bibliográficas.

Sin embargo, existen algunas carencias en el sistema ortográfico de la lengua española que imposibilitan la interpretación fonética de palabras del léxico propio si no están en nuestra competencia. Así, por la ortografía, *aguapié* y *rodapié* son fonológicamente ambiguas: /a.gua.'pie/ , /a.gua.pi.'e/ , /ro.da.'pie/ , /ro.da.pi.'e/ , la acentuación por sí sola no aclara si la pronunciación de la sílaba final ha de realizarse con diptongo o hiato. Esta anomalía del sistema evidencia que en la pronunciación de algunas palabras del español, además de las reglas generales, intervienen características idiosincrásicas, por lo que un sistema de transcripción fonética automática que quiera actuar con un margen de error virtualmente nulo ha de realizar un estudio exhaustivo de la fonética en el léxico, es decir, deberá dar cuenta de las propiedades fonéticas, tanto regulares, como particulares, de los elementos léxicos de la lengua; de ese modo, la representación resultante podrá ser considerada como un reflejo de la pronunciación. Explicitará:

- (1) el repertorio de fonemas;
- (2) cómo estos fonemas se agrupan en sucesiones silábicas según las estructuras admitidas en el sistema;
- (3) la sílaba marcada por el acento en las palabras tónicas;
- (4) los procesos fonológicos característicos de la lengua derivados del contacto de los sonidos y de su posición en la cadena hablada, teniendo en cuenta que los mismos fenómenos que se producen dentro de la palabra se producirán también en la concatenación de palabras si se dan las mismas condiciones -sinalefas, coalescencia, asimilaciones, resilabificación, reacentuación-;
- (5) las variaciones alofónicas (p.ej., distintas realizaciones en la neutralización de las oclusivas sordas en posición de coda silábica) y las alternancias (p.ej., fenómenos de yeísmo / lleísmo) que se den en la lengua.

El objetivo del Diccionario Electrónico Fonético del Español, proyecto que actualmente desarrollamos en el Laboratorio de Lingüística Informática del Departamento de Filología Española de la Universidad Autónoma de Barcelona², es crear un sistema de transcripción automática que dé cuenta de los problemas señalados. El diccionario fonético que proponemos constituirá una descripción del sistema fonológico de la lengua estándar que contendrá, tanto los fenómenos regulares, como los irregulares. En esta comunicación presentamos el tipo de información lingüística que consideramos necesaria para la consecución de nuestro objetivo.

1. El Diccionario Electrónico Fonético del Español

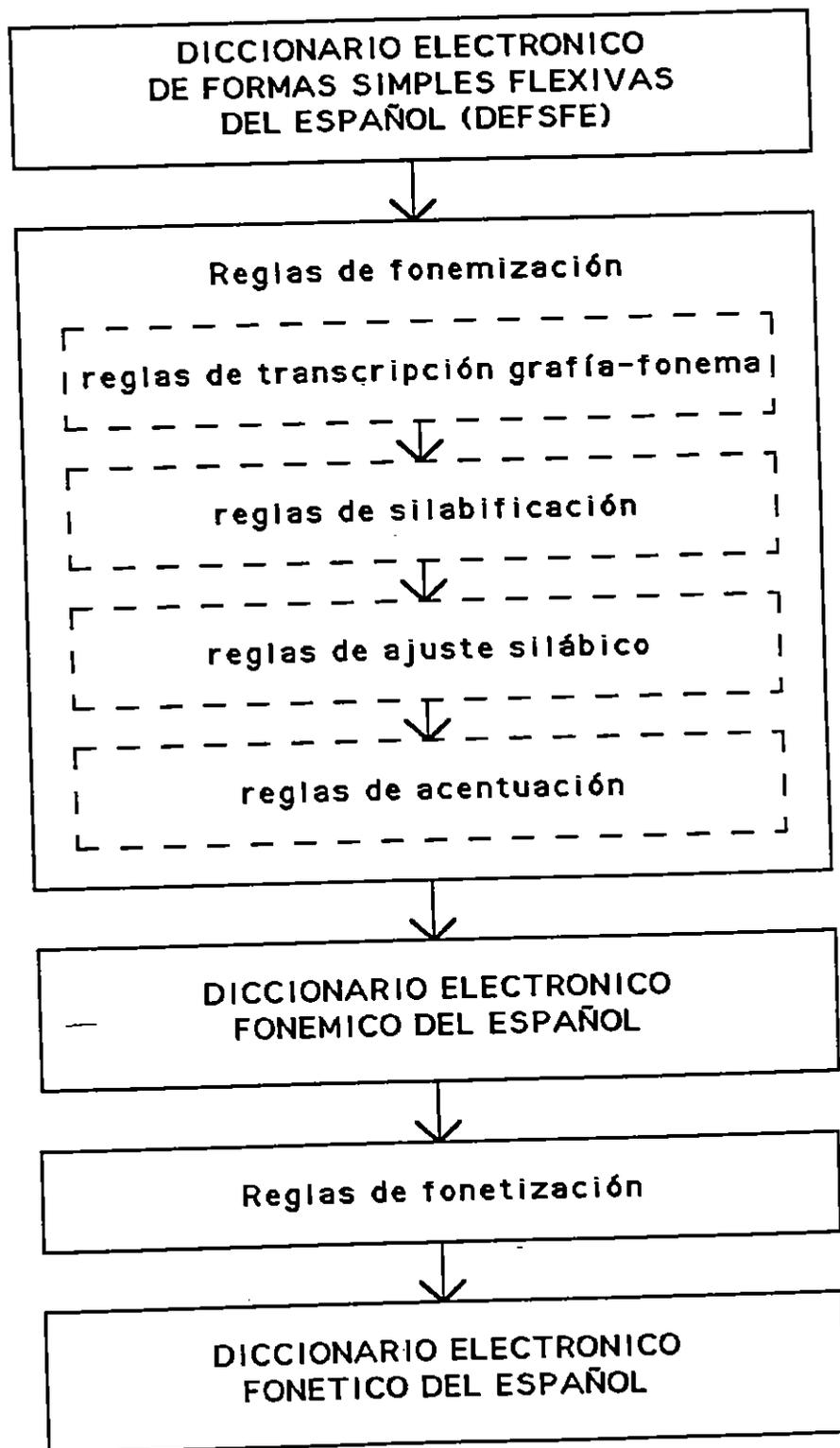
1.1. El Diccionario Electrónico Fonético y el Sistema de Diccionarios Electrónicos del Español

El Diccionario Electrónico Fonético del Español (DEFE) se integra en el Sistema de Diccionarios Electrónicos del Español (SDEE)³, que constituye una adaptación a nuestra lengua del sistema DELA, el sistema de diccionarios electrónicos desarrollado en el LADL (Laboratoire d'Automatique Documentaire et Linguistique de la Universidad Paris VII) [9]. El SDEE está formado por un «Diccionario Electrónico de Formas Simples del Español» (DEFSE), que consta de 70.000

² Proyecto financiado por el MEC (Beca de Formación del Personal Investigador. PN/90).

³ El SDEE es un proyecto de investigación que se está desarrollando en la Universidad Autónoma de Barcelona bajo la dirección del prof. C. Subirats, con la cooperación científica del Laboratoire d'Automatique Documentaire et Linguistique de la Universidad Paris VII y con financiación de la Dirección General de Telecomunicaciones (TIC 90-403).

Figura 1. Estructura del Diccionario Electrónico Fonético del Español (DEFE).



entradas codificadas alfanuméricamente por sus propiedades categoriales y morfológicas, y un paquete de programas de generación de formas flexivas, que generan automáticamente el "Diccionario Electrónico de Formas Simples Flexivas del Español" (DEFSFE). El DEFSFE contiene 557.000 formas con información morfológica y categorial, las cuales constituyen todas las formas posibles del léxico general de la lengua española. Del SDEE también formará parte el «Diccionario Electrónico de Formas Compuestas del Español», actualmente en fase de elaboración.

El DEFE consiste en una doble transcripción, fonémica y fonética, del «Diccionario Electrónico de Formas Simples Flexivas del Español». Está concebido como un sistema modular: el resultado de la aplicación de cada grupo de reglas de transcripción es la entrada para la aplicación de otro grupo, en una sucesión ordenada. Su estructura está representada en la Figura 1.

El DEFE se ha desarrollado adaptando al español el algoritmo creado por el profesor Eric Laporte [5] para el francés. La adaptación de este sistema se ha realizado teniendo en cuenta las necesidades de nuestra lengua y los objetivos de nuestra investigación, p.ej., hemos incluido los módulos de silabificación y acentuación, no contemplados para el francés. El sistema de fonetización de Laporte propone la generación de un diccionario electrónico fonético que parte de la fonemización de las formas de base y de su posterior fonetización al generar las formas flexivas, resuelve así las irregularidades que en francés se dan en la flexión. Este algoritmo se caracteriza por la aplicación global de las reglas y por la separación entre las reglas y el programa.

1.2. La Estructura del Diccionario Electrónico Fonético del Español

Sobre la representación ortográfica del DEFSFE actúa un paquete de «Reglas de fonemización», formado por cuatro subgrupos:

1. «Reglas de transcripción grafía-fonema», que asignan una representación en signos fonéticos a la cadena grafemática eliminando las irregularidades del sistema ortográfico español.
2. «Reglas de silabificación», que tienen en cuenta las restricciones del sistema fonológico español en las combinaciones de vocales y consonantes y la presencia del acento gráfico en la realización de hiatos.
3. Las reglas de silabificación realizan un corte silábico normativo, que resulta erróneo en los extranjerismos. El paquete de "Reglas de ajuste silábico" elimina aquel material gráfico que no tiene una interpretación fonética en nuestra lengua y realiza la resilabificación pertinente en cada secuencia (p.ej. /án-g-s-trom/ → /áns-trom/). Con este paquete se reduce el número de reglas de conversión grafía-fonema de baja productividad y permite la reducción de las listas de excepciones.
4. Las "Reglas de acentuación" atienden a la división silábica y a los finales de las palabras, asignando el acento fonético en las tónicas sin acento gráfico y descartan la acentuación de las átonas.

La salida de este primer grupo de reglas constituirá, de hecho, un primer diccionario, el "Diccionario Electrónico Fonémico del Español", una transcripción fonémica del léxico general de la lengua española que será una representación abstracta de lo común y distinto de la variación existente en la misma.

Sobre el "Diccionario Electrónico Fonémico del Español" actúa un último paquete de "Reglas de fonetización", que realizan una conversión fonema-sonido transcribiendo los distintos alófonos de los fonemas españoles y cuya salida constituirá el "Diccionario Electrónico Fonético del Español", una representación de la variación existente en la pronunciación de nuestra lengua.

1.3. Formalismo de las reglas

El conjunto de reglas que generan el DEFE son reglas contextuales de aplicación local. Tienen la particularidad de no ser ordenadas, es decir, dentro de cada uno de los paquetes se aplican todas simultáneamente. El sistema detecta de modo automático cuando dos o más reglas dan cuenta del mismo fenómeno, por lo que es necesario un estudio exhaustivo de los contextos exactos de aplicación de cada regla.

El formato de las reglas consta del foco sobre el que se aplica el cambio propuesto y tres contextos precedentes y cuatro siguientes opcionales. La ausencia de un signo en los campos contextuales indica que no hay restricciones en la aplicación de la regla. Los elementos pueden ser negados, median

Existen tres tipos de reglas:

(1) Reglas que modifican un signo en otro en una relación unívoca. Ej., la regla que transcribe el grafema *g* como [x] en el contexto precedente a las vocales *e, i*.

> g = x
 " " " g,, eéí,, " " "
 x0.

(2) Reglas que suprimen un elemento, es decir, reglas que realizan una conversión nula. Ej., la regla que elimina la *u* en el dígrafo *gu* ante las vocales *e, i*.

> u = 0
 " " g,, u,, eéí,, " " "

(3) Reglas que insertan un elemento, es decir, que transcriben un signo de entrada como dos en la salida. Ej., la regla que transcribe el grafema *x* como [ks].

> x = ks
 " " " x,, " " " "
 k0,s0.

Cada una de las reglas formulada como una hipótesis que verificamos mediante su aplicación en el sector del léxico que contiene el fenómeno estudiado.

2. La información lingüística en el DEFE

Para cumplir el objetivo del DEFE (la fonetización del léxico del español con un margen de error virtualmente nulo de modo que refleje la pronunciación normativa de la lengua) la información grafemática es insuficiente. A continuación indicamos el tipo de información lingüística que tenemos en cuenta en esta investigación, de la cual hay aspectos que no siempre han sido contemplados en los trabajos sobre la fonetización del español anteriores al nuestro:

1. Información fonológica:

1.1. Sonidos con una posición no permitida en la cadena, como la [s] de *standing* y la [f] de *frático*.

1.2. Restricciones en las combinaciones de grupos consonánticos y vocálicos en las agrupaciones silábicas y en los límites de las mismas.

1.3. Silabificación, considerando que el límite silábico -la realización de un hiato- es pertinente en casos como *pie* (nombre) /*pié* (verbo *piar*).

1.4. El acento, al ser el español una lengua de acento léxico libre.

1.5. Procesos fonológicos (neutralización, asimilación, coalescencia, diptongación ...).

2. Información morfofonológica.

La presencia de un diptongo o de un hiato es una propiedad idiosincrásica en determinados grupos de verbos, como en los de la 1ª conjugación cuya raíz acaba en [i], p.ej., *odiar* : [o.'djar] vs *[o.di.'ar] / *vaciar* : [ba.ʔi.'ar] vs *[ba.'ʔjar]. Esta propiedad se cumple en todo el paradigma verbal, por lo que puede ser marcada en los infinitivos del "Diccionario Electrónico de Formas Simples del Español" previamente a la generación de las formas flexivas; de otro modo, en la medida que no depende del contexto fónico, sería necesario un determinado número de reglas *ad hoc* o listas de excepciones para dar cuenta del fenómeno.

3. Información categorial.

3.1. La categoría de la palabra tiene consecuencias en la acentuación de la misma: existen clases de palabras átonas, p.ej., las conjunciones, y clases con doble acentuación, como los adverbios en *-mente*.

3.2. Todas las palabras acabadas en *-uar* forman diptongo cuando a este contexto precede una consonante velar (*jaguar* : [xa.'@war]). Este hecho es especialmente relevante en los verbos de la 1ª conjugación cuya raíz acaba en [u] ya que permite establecer dos subclases: los que diptongan por el contexto velar, p.ej., *averiguar* : [a.be.ri.'@war], y los que no poseen ese contexto, realizados todos con hiato, p.ej., *actuar* : [a@.tu.'ar]. Este fenómeno, que se reproduce en todo el paradigma de flexión de estos verbos [10], explicita un condicionamiento fonológico ligado a una determinada categoría.

3. Conclusiones

Hemos mostrado que existen fenómenos fonológicos idiosincrásicos en el léxico del español para los que no es suficiente la información grafemática cuando han de ser tratados en un sistema automático de transcripción: en el DEFE, para dar cuenta de la oposición combinatoria de las vocales en contacto «diptongo/hiato» de determinadas clases de verbos, necesitamos remitir a un estado previo a la fonetización, marcando esa idiosincrasia al generar la flexión verbal en el DEFSFE.

Ese hecho es coherente con el carácter ordenado, desde un punto lingüístico, con el que está concebido el diccionario fonético que proponemos. En él, cada grupo de reglas describe un aspecto distinto del sistema fonológico del español, en un doble plano de descripción fonémica (nivel segmental - combinaciones de segmentos en unidades fonológicas mayores (sílabas) - nivel suprasegmental del acento) y fonética (variaciones alofónicas). También aportamos un tratamiento automático para las irregularidades de los extranjerismos que reduce las tradicionales listas de excepciones y que los sistematiza en la estructura de nuestra propuesta. No obstante, la investigación que hemos presentado está en fase de desarrollo, por lo que los resultados son preliminares y posiblemente surgirán cuestiones no previstas inicialmente que deberán ser integradas en el diseño general.

Referencias bibliográficas

- [1] Lammens, J.M.G. (1987) «A Lexicon-Based Grapheme-to-Phoneme Conversion System», en Laver, J. & Jack, M.A. (eds), *European Conference on Speech Technology. Edinburgh, September, 1987* : 281-285.
- [2] *Journal of Phonetics*, 18,3. Theme issue: *Phonetic Representation*. Mary E. Beckman (ed.). July 1990.
- [3] Sobre la fonetización automática del español:
Cabrera, C. , Contini, M. & Boë (1991) «La Phonétisation du Castillan», *Actes du XII^{ème} Congrès International des Sciences Phonétiques*, Aix-en-Provence, août 1991. Vol. 4. Université de Provence : 114-117.

- Martí, J. & Niñerola, D. (1987) «SINCAS: un conversor texto-voz en castellano. *Procesamiento del lenguaje natural*. *Boletín SEPLN*, nº 5 : 112-122.
- Martínez, M. , Pardo, J.M. , Borrajo, A. , Santos, A. , Barrio, C.L. , Muñoz, E. & Qullis, A. (1986) "Conversión automática texto-habla y su relación con el lenguaje natural", en C. Martín Vide (ed.), *Lenguajes naturales y lenguajes formales*, vol. II. Barcelona, P.P.U. : 197-209.
- «Monográfico sobre reconocimiento y síntesis del habla», *Boletín SEPLN*, nº 6, junio, 1988.
- Santos, J.M. & Nombela, J.R. (1982) «Text-to-Speech conversion in Spanish, a complete rule-based system». *Proceedings IEEE*.
- Ueda, H. (1986) «Transcripción automática del texto español», *Lingüística Española Actual*, IX,2 : 217-240.
- [4] Mosterín, J. (1981) *La ortografía fonémica del español*, Alianza, Madrid.
- [5] Laporte, E. 1988 *Méthodes algorithmiques et lexicales de phonétisation de textes. Applications au français*. Tesis doctoral, Université Paris VII, LADL.
- [6] Manuales de pronunciación española.
- Bello, A. (1955) *Principios de ortología y métrica de la lengua castellana*, en *Obras Completas I*. Caracas.
- Canellada, M.J. & Kuhlmann Madsen, J. (1987) *Pronunciación del español. Lengua hablada y literaria*, Castalia, Madrid.
- Navarro Tomás, T. (1919) *Manual de pronunciación española.*, Madrid, CSIC. 24ª edición, 1982.
- [7] Santiago, R. (1989) «Hacia la reconstrucción histórica de un problema de adaptación gráfica y fonética: La normativa sobre W», *Boletín de la Real Academia Española*, tomo LXIX : 41-63.
- [8] Lehiste, I. (1988) *Lectures on Language Contact*. Cambridge, Mass. The MIT Press.
- [9] Subirats, C. (1989) "Verbal morphology in the Electronic Dictionary of Spanish", *Lingvisticae Investigationes*, XIII.1 : 179-201.
- Subirats, C. (1992) "Verbal, nominal and adjectival inflection in the Electronic Dictionary of Simple Forms of Spanish", *Lingvisticae Investigationes*, XVI.2 (en prensa).
- [10] Seco, M. (1986⁹) *Diccionario de dudas y dificultades de la lengua española*. Espasa Calpe, Madrid. (cf. sv. «actuar»).

