

Conversión Fonética Automática con Información Fonológica para el Gallego

Automatic Phonetic Conversion with Phonological Knowledge for Galician

Marcos Garcia

Centro de Investigación en
Tecnoloxías da Información (CITIUS)
Univ. de Santiago de Compostela
marcos.garcia.gonzalez@usc.es

Isaac González López

Cilenis Language Technology
Santiago de Compostela
isaacjgonzalez@cilenis.com

Resumen: En este artículo se presenta un sistema de conversión fonética con información fonológica para el gallego. El conversor permite evaluar diferentes procesos fonológicos como la silabificación o la derivación fonológica, así como transcribir automáticamente léxicos con información fonética de diferentes variedades lingüísticas. Varios módulos componen el sistema presentado: (i) conversión fonológica, que genera una representación fonológica a partir de un *input* ortográfico; (ii) silabificación, (iii) asignación de acento y (iv) conversión fonética, que produce diferentes representaciones fonéticas (de diversos dialectos) desde una representación fonológica. La precisión del sistema, evaluado en un *corpus* periodístico, supera el 99 %.

Palabras clave: conversor fonético, fonología, gallego, portugués.

Abstract: This paper presents a phonetic conversion system with phonological knowledge for Galician. The system allows us to evaluate several phonological processes, like syllabification or phonological derivation, as well as to transcribe lexicons with phonetic annotation of different linguistic varieties automatically. The system consists of the following modules: (i) phonological conversion, which generates an underlying representation from an orthographic input; (ii) syllabification, (iii) stress marker and (iv) phonetic conversion, which produces different phonetic representations (of various dialects) from a single phonological input. The system has been evaluated in a journalistic corpus, achieving a precision higher than 99 %.

Keywords: phonetic converter, phonology, Galician, Portuguese.

1. *Introducción*

Los sistemas de transcripción fonética automática (*grafema a fono*) se utilizan con diferentes finalidades en una gran variedad de aplicaciones. Estas herramientas destacan por su utilización en sistemas de síntesis de voz, donde convierten un texto en su representación fonética, posteriormente transformada en señales acústicas.

Además de esta finalidad, los sistemas *grafema a fono* también son utilizados por lingüistas para generar automáticamente transcripciones fonéticas de textos de diferentes tipologías. Estas transcripciones, normalmente realizadas de acuerdo con una variedad estándar, se usan como punto de partida para la corrección manual por parte del profesional. Así mismo, los conversores *grafema a fono* son de gran utilidad para el aprendizaje de lenguas extranjeras, permitiendo conocer con diferen-

te grado de detalle las posibles realizaciones fonéticas de una palabra o texto.

De modo general, los conversores fonéticos realizan una transcripción directa entre grafemas y segmentos fonéticos (<campo>→[ˈkaŋ.po]), por lo que la información fonológica que contienen las representaciones es normalmente limitada. En este sentido, una mayor presencia de información de carácter fonológico (que contenga diferentes niveles de representación, por ejemplo) introduce algunas novedades en estas aplicaciones.

Así, a través de una representación fonológica intermedia (<campo>→/kanpo/), un sistema de estas características permitirá que el proceso de conversión entre la representación fonológica y la fonética contenga reglas de mayor abstracción lingüística (que incluyan información fonológica y no ortográfica). De este modo, con variaciones en estas reglas,

es posible generar diferentes formas fonéticas de dialectos que compartan una misma representación fonológica (/kanpu/→[^hkam.po] o [^hkãj.pu]).

En este trabajo presentamos una aplicación que implementa teorías lingüísticas para la conversión fonológica y fonética. El sistema se compone de los siguientes módulos: (i) conversión entre una representación ortográfica y una representación fonológica, (ii) silabificación, (iii) asignación de acento y (iv) derivación entre la representación fonológica y la representación fonética.

Este último módulo, de conversión fonética, permite añadir, eliminar y modificar reglas o restricciones fonético-fonológicas, para su validación tanto a pequeña como a gran escala. Las reglas, añadidas como parámetros, se pueden activar o desactivar en cada ejecución. Esto permite, por un lado, obtener automáticamente léxicos con información fonética de diferentes variedades lingüísticas, ejecutando el conversor con los parámetros seleccionados. Por otro lado, facilita la evaluación de varios aspectos de la fonología teórica, introduciendo nuevas reglas lingüísticas y verificando su funcionamiento en la transcripción de grandes cantidades de *corpora*

Además de permitir utilizar la ortografía de mayor uso en Galicia (Instituto da Lingua Galega y Real Academia Galega, 2005), el sistema contiene un módulo de conversión fonológica desde el sistema ortográfico portugués, por lo que puede ser utilizado, con resultados similares, desde texto escrito con esta ortografía (también utilizada para representar ortográficamente el gallego).

El sistema, distribuido con licencia libre GPLv3,¹ implementa modelos fonológicos del gallego y produce una salida con alfabeto SAMPA² (con conversión opcional a IPA (Ladefoged, 1990)) con valores de precisión superiores al 99%.

Tras esta sección introductoria, la sección 2 presenta sucintamente el trabajo relacionado. La sección 3 describe los diferentes módulos del sistema. En la sección 4 realizamos una evaluación en texto real y, finalmente, la sección 5 concluye este trabajo.

2. Estado del arte

En los últimos años han sido publicados diversos trabajos que describen varios aspectos

de la conversión fonética automática para gallego y para portugués. Así, Braga y Coelho (2006) y Braga y Freixeiro (2007) han implementado algoritmos de conversión grafema-fono para gallego (y portugués), con resultados superiores al 98% de precisión.

La Universidade de Vigo, junto con el Centro Ramón Piñeiro para a Investigación en Humanidades, han desarrollado *Cotovia*,³ un sistema TTS (*text-to-speech*) que contiene módulos de conversión fonética. Dentro de este proyecto, González González et al. (2008) describen las principales características a tener en cuenta en la elaboración de un sistema de conversión texto a fono para el gallego, como la asignación de la abertura de las vocales tónicas, la transcripción de la letra <x> (que puede corresponderse con diferentes segmentos fonológicos), entre otras características segmentales y suprasegmentales.

Etapas posteriores, como la desambiguación de homógrafos heterófonos son abordadas por Mourín et al. (2009). Así mismo, Campillo et al. (2010) describen el proceso de elaboración de un *corpus* y un léxico dirigidos a sistemas de síntesis de voz en gallego.

Estas aplicaciones se centran en la transcripción fonética para aplicaciones de conversión texto a voz, por lo que la información fonológica utilizada es limitada y adaptada a los objetivos de este tipo de herramientas.

En portugués existe un mayor número de trabajos que tratan diferentes procesos de la conversión fonética (Teixeira, Oliveira, y Moutinho, 2006; Braga, Coelho, y Resende Jr., 2006; Paulo et al., 2008; Siravenha et al., 2008, entre otros). De modo general, estas implementaciones también se destinan a módulos de sistemas TTS. Con todo, Oliveira, Moutinho, y Teixeira (2005) comparan, en el proceso de división silábica un sistema basado en la aplicación de teorías fonológicas con un conversor directo grafema→fono, con mejores resultados para el primero.

Seara et al. (2003) incluyen procesamiento fonológico en la generación de variantes fonéticas destinadas a sistemas de reconocimiento de voz para diferentes dialectos.

Por último, Vigário, Martins, y Frota (2006) describen FreP, *software* que genera representaciones con información fonológica de grandes cantidades de *corpora*. FreP se destina a la obtención de información sobre

¹<http://www.gnu.org/licenses/gpl.txt>

²<http://www.phon.ucl.ac.uk/home/sampa/index.html>

³<http://www.gts.tsc.uvigo.es/cotovia>

las frecuencias y contextos de aparición de diferentes clases y segmentos fonológicos.

3. Arquitectura del sistema

El sistema presentado en este trabajo realiza, por palabras individuales, la conversión entre las siguientes representaciones: *ortografía* → *fonología* → *fonética*.

El primer módulo realiza la conversión de una palabra ortográfica a una representación fonológica. A partir de ahí, se aplica la teoría silábica ataque-rima (Blevins, 1995, por ejemplo) para construir los elementos silábicos y sus constituyentes. Un siguiente módulo asigna el acento, aprovechando la información ortográfica para encontrar la sílaba tónica. La derivación fonológica (Mohanan, 1986) continúa con la aplicación de un conjunto de reglas léxicas y posléxicas, que dan lugar a la representación fonética de la palabra analizada. La Tabla 1 ejemplifica este proceso.

La transcripción de unidades oracionales es posteriormente realizada a través de módulos específicos, no tratados en el presente trabajo.

3.1. Conversión fonológica

El primer proceso realizado por el sistema consiste en convertir, dada una palabra aislada, la representación ortográfica en una representación fonológica. Como hemos dicho, este proceso puede ser realizado desde dos ortografías, por lo que se divide en dos módulos diferenciados.

3.1.1. Ortografía ILG/RAG

El módulo de conversión fonológica (con ortografía ILG/RAG) se compone de una serie de reglas transformacionales, unidas a un conjunto de listas de excepciones que contienen formas cuya representación no es deducible a través de la ortografía.

Las reglas substituyen los caracteres ortográficos por representaciones fonológicas (salvo en los casos en que coinciden: por ejemplo: - /b/). La Tabla 2 muestra un ejemplo de este tipo de transformaciones. Además de pequeños conjuntos de excepciones, como las formas <unha> y derivadas, o formas como <abrogar> (donde <r> representa /r/ y no /r/), la principal lista de excepciones utilizada en esta etapa contiene palabras con <x> cuya representación no es deducible por el contexto. Este grafema puede ser representado fonológicamente como una sibilante /ʃ/ (<caixa>), o como /ks/ (<sexo>); por ser

<viño> / <vinho>	Repr. Ortográfica
/bipno/	Repr. Fonológica
/bi.ɲo/	<i>División Silábica</i>
/'bi.ɲo/	<i>Asignación de Acento</i>
/'bĩ.ɲo/	<i>Nasalización</i>
/'bĩ.ɲʊ/	<i>Vocalismo Átono</i>
['bĩ.ɲʊ]	Repr. Fonética

Tabla 1: Arquitectura del sistema, ejemplificado con la palabra <viño> (ILG/RAG) o <vinho> (ortografía portuguesa).

Regla	Entrada	Salida
$v \rightarrow /b/$	vaca	/b/aca
$gu(ei) \rightarrow /g/(ei)$	guerra	/g/erra
$qu(ei) \rightarrow /q/(ei)$	quente	/k/ente
$\ddot{u}(ei) \rightarrow /u/(ei)$	fragüe	frag/u/e
$ch \rightarrow /tʃ/$	chamar	/tʃ/amar
$c[aou] \rightarrow /k/[aou]$	carro	/k/arro
$ll \rightarrow /ʎ/$	mollo	mo/ʎ/o
$rr \rightarrow /r/$	carro	ca/r/o
$(nmls)r \rightarrow /r/$	enredar	en/r/edar
$x(\text{conson.}) \rightarrow /ks/$	extra	e/ks/tra
$x \rightarrow /ʃ/$	xeito	/ʃ/eito

Tabla 2: Selección de reglas de conversión fonológica (ortografía ILG/RAG).

menos frecuentes, las palabras incluidas en la lista de excepciones son las que contienen la estructura /ks/. La lista no contiene formas cuya representación es deducible y que, por lo tanto, son analizadas por las reglas generales: <x> seguido de consonante o en posición final, por ejemplo.

Otra de las dificultades de la transcripción fonética del gallego (y del portugués) consiste en la conversión de las vocales <e> y <o> (sobre todo en posiciones tónicas y pretónicas) por sus equivalentes fonológicos, que pueden ser /e, ε/ y /o, ɔ/, respectivamente. Aunque existen contextos favorables para la aparición de uno u otro segmento fonológico, en un gran número de casos no es predecible. A esto hay que sumarle el hecho de que el gallego no tiene un estándar fonético dominante, y que dialectalmente existen variaciones en el timbre de las vocales.

Entre los casos problemáticos de asignación de timbre vocálico tónico encontramos formas nominales (“c/e/sta” vs “f/ε/sta”) y verbales (“c/o/mo” vs “c/ɔ/me”). Además,

existen vocales abiertas en diferentes posiciones de formas verbales (tónica, pretónica, etc.), que normalmente se corresponden con la vocal radical y cuyo timbre se mantiene en toda la conjugación (y no sólo en posición tónica): “/ε/levar, /ε/leva, /ε/levando”, etc.

Téngase en cuenta que, además de estos casos, existen formas cuya representación fonológica depende de prefijos concretos, que pueden incluir hasta tres vocales /ε, ɔ/ (tónica, pretónica, antepretónica, etc.): “c/ɔɔ/perante”, “/ε/ctrocutar”, etc. Este tipo de formas son analizadas por nuestro sistema a través de un conjunto de reglas específicas que contienen este tipo de prefijos.

Salvo estos casos, la asignación del timbre vocálico es realizado a través de listas de palabras con vocales tónica y/o pretónica abiertas (/ε, ɔ/), unidas a reglas simples de lematización. Las listas utilizadas en las evaluaciones de este trabajo tienen como referente a Regueira (2010) para las formas nominales, e ILG/RAG (2005) para las formas verbales (extraídas automáticamente de grandes léxicos). Además, ha sido incluido un conjunto de nombres propios frecuentes (países, capitales, ayuntamientos gallegos, así como nombres de pila frecuentes) que tienen /ε, ɔ/. Estas listas de excepciones tienen forma ortográfica por lo que no es necesaria la transcripción fonética de las palabras para ampliarlas.

3.1.2. Ortografía portuguesa

El módulo de conversión fonológica desde ortografía portuguesa contiene, del mismo modo, reglas universales y listas de excepciones. El sistema ortográfico portugués tiene otro tipo de ambigüedades, que son tenidas en cuenta en este proceso.

Las formas con <qu> y <gu> seguidas de <e> (y de <i>), pueden ser representadas (en función de la palabra) como /ke/, /ge/, /kue/ o /gue/.⁴ Esta ambigüedad es resuelta gracias a listas de palabras (extraídas automáticamente de diccionarios paralelos de portugués brasileño y europeo) cuya representación contiene /u/.

Las formas terminadas con un diptongo nasal en ortografía portuguesa (<camião>, <pão>), así como los plurales en -ães, -ões) pueden representar, en gallego, una secuencia de vocal /o/ o /a/ seguida de un segmento nasal: “cami/on/”, “p/an/”. En este caso, la

⁴O con /kw/, /gw/, en función de la propuesta teórica utilizada.

lista de excepciones contiene las formas representadas como /an/, debido a su menor frecuencia. De modo similar a las anteriores, estas listas han sido automáticamente extraídas de léxicos paralelos entre portugués europeo y gallego.

El léxico que contiene las palabras con /ε/ y /ɔ/ está escrito en ortografía ILG/RAG, por lo que un *lookup* desde formas introducidas en ortografía portuguesa no tendría amplia cobertura. Para eso, este módulo contiene un pequeño conjunto de reglas de adaptación ortográfica, que convierten una palabra escrita en portugués en ortografía ILG/RAG (-ção→-ción, lh→ll, -m→-n, j/g→x, nh→ñ, adaptaciones de acento, etc.). Este método, a pesar de no ser totalmente efectivo⁵ ya ha sido probado con éxito, entre otros, en sistemas de traducción automática (Malvar et al., 2010). Además de esto, la ortografía portuguesa marca, en algunos casos, el timbre de la vocal tónica (<ó> vs <ô> o <é> vs <ê>, etc., donde el acento circunflejo indica vocal cerrada), por lo que muchas formas no necesitan ser buscadas en listas de excepciones.

Además de estas transformaciones orientadas a la realización de *matching* con las listas de excepciones, en la conversión de las formas escritas con ortografía portuguesa existen otras formas que son comunes en Portugal, pero cuya presencia en Galicia es escasa: formas verbales irregulares como <fiz>, <quis>, etc., u otras palabras como <ontem> o <sim>, y sufijos como <agem> son adaptadas a las formas gallegas más frecuentes (<fixem>, <quixem>, <onte>, <si>, <axe>).

En síntesis, el módulo de conversión fonológica toma como entrada una palabra (escrita en ortografía portuguesa o ILG/RAG) y produce una representación fonológica que incluye, entre otras características, el timbre de las vocales representadas como <e> y <o>.

3.2. Silabificación

Obtenida la forma fonológica de la palabra, el siguiente proceso consiste en construir una representación de cada palabra que incluya las sílabas (σ) y sus constituyentes internos: ataque (A) y rima (R), que a su vez se compone de núcleo (Nc) y coda (Cd) (Fi-

⁵Nótese que una forma introducida con ortografía portuguesa como <suporte> no será encontrada en la lista, ya que en ILG/RAG la palabra se escribe <soporte>, y la conversión u→o no es universal.

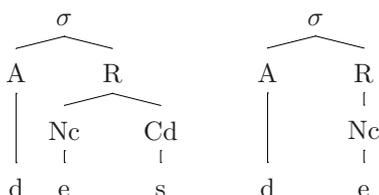


Figura 1: Ejemplo de la estructura silábica ataque-rima de la palabra <desde>.

gura 1).

Para esto, implementamos una variante de la teoría silábica ataque-rima (propuesta para portugués en Mateus y Andrade (2000) e utilizada en Oliveira, Moutinho, y Teixeira (2005)), con algunas simplificaciones.

La propuesta original utiliza una escala de sonoridad que atribuye a cada clase de segmentos —oclusivas, laterales, vocales, etc.— un valor de sonoridad (Selkirk, 1984). A través de este valor, los segmentos ocuparán una u otra posición en la estructura silábica, que a su vez es dependiente de la lengua. Así, en gallego el núcleo es obligatorio y sólo podrá ser ocupado por segmentos vocálicos, el ataque es opcional y puede ser simple (/t_{ona}/) o ramificado (/t_rona/), etc.

El sistema aquí presentado conoce de antemano qué posiciones puede ocupar cada segmento en la estructura silábica, por lo que no utiliza la escala de sonoridad. Así, el algoritmo de silabificación es el siguiente:

1. Se marcan como núcleos los diptongos decrecientes (/kaifa/).
2. Se marcan como núcleos las vocales (/kasa/).
3. Se marcan los ataques /tʃ/ y los complejos (segmentos oclusivos seguidos de líquidas: /klabe/, /prato/, etc.). Nótese que excepciones como /abrogar/ o /subrutina/ no son aquí analizadas, ya que la secuencia
 ha sido transformada en /br/ durante la conversión fonológica.
4. Se marcan los ataques simples seguidos de un núcleo ya identificado (/kasa/).
5. Se marcan las codas complejas y no patrimoniales (/instituto/, /abstrakto/).
6. Se marcan las codas simples (/mar/).
7. Se aplican algunas excepciones con codas infrecuentes (extranjerismos o cultismos): /ganster/, /persikaθia/, etc.

Este algoritmo tiene pequeñas variaciones en función de la ortografía escogida como *input*; así, el encuentro vocálico /ui/ no se considera diptongo con ortografía ILG/RAG (salvo excepciones como <uitar>). Por el contrario, con ortografía portuguesa se considera diptongo (salvo excepciones: <ruim>). Téngase en cuenta que variantes como <moito–muito>, <coidar–cuidar>, etc. coexisten en gallego, aunque la primera es la forma preferente en ortografía ILG/RAG, mientras que la segunda es la forma habitual en ortografía portuguesa. Con respecto a los diptongos crecientes (que pueden interpretarse como *no fonológicos*: “soc/ial/” → “soc[jal]”), estos pueden ser transformados en el proceso de silabificación, o como una regla más de la derivación fonológica.

Aplicado el algoritmo de silabificación, el sistema permite incluir reglas optativas para el tratamiento de determinadas secuencias. Así, una regla parametrizable transforma una secuencia /psi/ (<psicólogo>) en /si/ o /pe-si/ (eliminando una sílaba sin núcleo, o incluyendo una vocal en un núcleo vacío).

Una vez construida la representación silábica, el sistema marca las fronteras entre estos constituyentes (i) antes de cada ataque, (ii) antes de un núcleo en posición inicial, (iii) entre dos núcleos y, finalmente, (iv) entre una coda y un núcleo.

3.3. Asignación de acento

La función del siguiente módulo del sistema consiste en la detección de la sílaba tónica de la palabra analizada. Una vez que el sistema mantiene la ortografía de la palabra, este proceso es relativamente simple. La asignación de acento varía en función de la ortografía, ya que las reglas de acentuación de las dos ortografías son diferentes.

Básicamente, este módulo verifica primero si la palabra introducida tiene alguna vocal acentuada gráficamente (si la tiene, marca la sílaba que la contenga como tónica). Si no la tiene, aplica inversamente las reglas de acentuación de cada ortografía hasta detectar la sílaba tónica.

Finalmente, existe una lista de palabras que no tienen acento (preposiciones y pronombres átonos: <con>, <que>, <en>, <te>, <me>, etc.). Algunas de ellas, como <o> y <a>, son marcadas como homógrafas, porque pueden ser o no acentuadas en

función de su categoría morfosintáctica.

3.4. Derivación fonológica

El módulo de derivación fonológica (independiente ya de la ortografía seleccionada), aplica una serie de reglas transformacionales sobre la estructura fonológica producida. Las reglas se aplican secuencialmente, por lo que la salida de una será la entrada de la regla posterior. Así —del mismo modo que en la fonología léxica— el orden de aplicación de las reglas es fundamental en este proceso.

Actualmente, el sistema contiene un conjunto de reglas universales y otro de reglas opcionales (propias de una variedad dialectal o social concreta, etc.).

Entre las reglas universales se encuentran:

- Semivocalización: transforma determinados encuentros entre vocales de diferentes sílabas adyacentes en diptongos: /su.per.fi.θi.al/ → /su.per.fi.θjal/, salvo excepciones como /tri.al/, /di.al/ o /si.al/.
- Aproximantes: convierte los segmentos /b, d, g/ en aproximantes ([β̞, ð̞, ɣ̞]) después de vocal o coda [s, r, l].
- Vocalismo átono: cierra y centraliza las vocales átonas en posición final: /a/ → [ə], /e, i/ → [ɪ] y /o, u/ → [ʊ]. Esta regla tiene varios parámetros opcionales para representar diferentes grados de vocalismo átono (inicial, antepretónico, postónico medial, etc.).
- Nasalización: esta regla, también parametrizable, permite seleccionar el tipo de segmento nasal implusivo (velar, compartiendo la articulación de la consonante siguiente, etc.), su posición en la estructura silábica (coda o núcleo) o los contextos en los que las vocales también son nasalizadas (inicio absoluto, entre dos consonantes nasales, etc.).

Las reglas opcionales —de dialectos o registros concretos— implementadas son:

- Zetacismo: el conversor incorpora dos modelos de sibilantes: (i) con zetacismo, que representa el mismo sistema que el castellano estándar (/s, θ, f/, con un segmento más: /f/) y (ii) sin zetacismo (que aplica el seseo): simplificación de sibilantes /s, θ/ en /s/. Actualmente, y a diferencia de las restantes reglas, el conversor

interpreta la selección de este parámetro como fonológica: así, el modelo con zetacismo incluye /θ/ y [θ] en la representación fonológica y fonética, mientras que, sin zetacismo, obtendremos /s/ y [s], respectivamente. Así mismo, la realización del segmento fricativo coronal /s/ (ápicoalveolar, predorsodental, etc.) es también parametrizable.

- Codas complejas: esta regla simplifica las codas complejas en codas simples (p. ej. /bs/ → /s/), y elimina codas no patrimoniales (p. ej. /k/) una vez que no son frecuentes en un gran número de hablantes y registros gallegos.
- *Gheada*: el segmento fonológico /g/ en ataque es producido como [h] y, después de una consonante nasal, [k].
- Rotacismo de /s/: esta regla representa un segmento /s/ en coda en determinados contextos como [r]: /desde/ → [ˈder.dɪ]. Si no se aplica, la fricativa es representada como [z].

Las reglas de la derivación fonológica son similares a las reglas transformacionales de la fonología SPE (*The Sound Pattern of English*, Chomsky y Halle (1968)), pero aprovechan toda la información presente en el sistema, incluida la forma ortográfica. Además, pueden contener expresiones regulares, por lo que la introducción y modificación de reglas es un proceso simple.

3.5. Homógrafas heterófonas

Además de las palabras que pueden ser tónicas o no en función de su categoría morfosintáctica (referidas en 3.3), han sido recopilados dos conjuntos de formas homógrafas heterófonas (uno para cada ortografía). La representación fonológica (y fonética) de estas palabras varía (i) en función de su categoría (<elixir>: verbo “eli/f/ir” o sustantivo “eli/k.s/ir”), (ii) de su significado (<sede>: *lugar*, “s/ε/de” o *falta de agua*, “s/e/de”) o del modo verbal (<come>, imperativo, “c/o/me” o presente de indicativo, “c/ɔ/me”). Si la entrada tiene ortografía portuguesa, existen otros casos, como las heterófonas acabadas en -ão (<papão>, “pap/o/n” o “pap/a/n”) o verbos irregulares (<quis>, <quixem/quixo>).

Las listas de homógrafas heterófonas contienen unos 540 pares de palabras cada una (la portuguesa tiene 10 pares más), siendo

la mitad de ellos pares de formas verbales “imperativo-presente”. El sistema presentado en este trabajo realiza la transcripción en palabras aisladas por lo que en estos casos produce las dos formas posibles (y no desambiguadas por el contexto).

Por último, es preciso referir que el sistema no contiene todavía reglas específicas de tratamiento de extranjerismos. Con todo, han sido incluidas algunas reglas básicas, y excepciones para tratar las formas más frecuentes.

4. Evaluación

Para evaluar el funcionamiento del conversor, hemos utilizado un *corpus* periodístico (extraído del periódico electrónico Vieiros) de unas 5,600 palabras (1550 únicas), con un total de 31,200 caracteres. El *corpus* está escrito en ortografía ILG/RAG, y ha sido transcrito fonéticamente (en alfabeto SAMPA) y revisado manualmente. Salvo pequeñas variaciones de transcripción, los parámetros de representación/ejecución se corresponden con la propuesta fonética de Regueira (2010). Para desambiguar las formas homógrafas heterófonas de diferente categoría morfosintáctica (que el sistema transcribe de los dos modos posibles), se han aplicado un *tokenizador*, un lematizador y un etiquetador morfosintáctico libres (Carreras et al., 2004; Garcia y Gama-lo, 2010). Además, ya que no se ha utilizado un reconocedor de expresiones numerales y cantidades, este tipo de formas no se han tenido en cuenta en la evaluación (93 casos).

La Tabla 3 muestra los resultados de tres evaluaciones sobre el mismo *corpus*: *Simple*, que ignora los nombres propios y las siglas (ya que el sistema no las expande ni contiene un módulo específico de análisis de nombres propios). *No-Sig*, que evalúa los nombres propios, pero no las siglas y *All*, que considera tanto los nombres propios como las siglas. El módulo de asignación de acento obtuvo un 100 % de precisión.

La inclusión de las siglas y los nombres propios en la evaluación (*All*) reduce la precisión un 0,4 %. Si tenemos en cuenta que el análisis de las siglas no se ha implantado (la precisión de estas formas fue de un 70 %), y que los léxicos sólo incluyen algunos nombres propios frecuentes (precisión del 92 %), los resultados obtenidos por la evaluación *Simple* superan el 99,7 % de precisión, con 69 de 27.148 caracteres convertidos erróneamente.

Analizando con detalle esta evaluación

Eval	Car.	Corr.	Inc.	Prec.
<i>Simple</i>	27.148	27.079	69	99,75 %
<i>No-Sig</i>	31.057	30.923	148	99,52 %
<i>All</i>	31.243	31.069	203	99,35 %

Tabla 3: Resultados de tres evaluaciones sobre el mismo *corpus*. *Car.* indica el número total de caracteres, *Corr.* e *Inc.* las conversiones correctas e incorrectas, respectivamente, y *Prec.* la precisión (correctos/total).

Tipo	Tok	Total
<e/o>	15	20
<e/o> Homógr.	2	2
<j>	1	1
Diptongo/Hiato	37	46
Total	54	69

Tabla 4: Tipos de errores producidos. *Tok.* muestra el número de *tokens* con algún carácter mal convertido, y *Total* el número total de errores en el *corpus* (evaluación *Simple*).

(*Simple*), podemos observar los tipos de errores producidos por el sistema. En la Tabla 4 vemos como la asignación del timbre de las vocales <e> y <o> ha producido 20 errores (en 15 palabras diferentes). A estos errores hay que añadir dos formas homógrafas heterófonas que el etiquetador morfosintáctico analizó incorrectamente (<posto> y <termos>). Otro error ha sido producido por un extranjerismo (<judicialista>), cuya <j> no fue convertida correctamente.

Aparte de esto, el sistema produjo 46 errores (con 37 *tokens* diferentes) en el proceso de división silábica. Algunos hiatos fueron representados como diptongos (<reunir>: *[rew.'nir]) y otros diptongos como hiatos (<necesariamente>: *[ne.θe.sa.ri.a.'mẽj.ti]).

Teniendo en cuenta que los tres primeros casos se corresponden con la conversión de la representación ortográfica a la representación fonológica y que el último forma parte de la derivación fonológica, la Tabla 5 muestra la precisión de cada uno de estos módulos por separado, así como la precisión final *Ortografía*→*Fonética* (de nuevo, con la evaluación *Simple*).

Varios de los errores producidos durante la conversión de los grafemas <e> y <o> se deben que el sistema analizó algunas formas

Conversión	Err	Prec
<i>Ortografía</i> → <i>Fonología</i>	23	99,91 %
<i>Fonología</i> → <i>Fonética</i>	46	99,83 %
<i>Ortografía</i> → <i>Fonética</i>	69	99.75 %

Tabla 5: Precisión de la conversión *Ortografía*→*Fonología*, *Fonología*→*Fonética* y *Ortografía*→*Fonética* (evaluación *Simple*).

nominales (con /e, o/) como formas de verbos con la vocal radical abierta (/ε, ɔ/). Además de estos casos, cuya análisis necesita mayores léxicos o reglas de lematización verbal más sofisticadas, la gran parte de los errores aparecen en encuentros vocálicos (diptongos e hiatos). Así, es necesaria una revisión de este tipo de procesos con el objetivo de continuar mejorando la precisión del sistema.

5. Conclusiones y trabajo futuro

El presente trabajo describe la implementación de diversas teorías fonológicas (teoría silábica ataque-rima, fonología léxica) en la construcción de un conversor fonético con información fonológica para el gallego. El sistema, distribuido libremente bajo licencia GPLv3, permite utilizar dos ortografías diferentes (ILG/RAG y ortografía portuguesa), y emula (con algunas variaciones) el funcionamiento de un sistema fonológico basado en las teorías referidas. De este modo, el conversor genera la representación y derivación fonológicas, así como la representación fonética. Además, posibilita la adición, modificación y eliminación de reglas, cuya aplicación es, así mismo, opcional.

Estas características permiten transcribir automáticamente textos de dialectos o registros concretos, evaluar el funcionamiento de teorías fonológicas sobre grandes cantidades de *corpora*, o crear automáticamente léxicos con diferentes transcripciones fonéticas en función de los parámetros aplicados.

La evaluación sobre un *corpus* de texto real muestra que la precisión del sistema se sitúa en el estado del arte, a pesar de que, principalmente en el análisis de extranjerismos y nombres propios, es necesario un trabajo de perfeccionamiento. Además, son también necesarias evaluaciones con *corpora* de mayor tamaño, con diferente *input* ortográfico y aplicando reglas de derivación fonológica de diferentes registros y dialectos gallegos.

En este sentido, el trabajo futuro se centra (i) en la depuración de errores (principalmente de tratamiento de diptongos e hiatos, y en la lematización verbal), (ii) en la inclusión de reglas de análisis para la adaptación de extranjerismos, (iii) en la introducción de algunas reglas de asignación de timbre vocálico para palabras desconocidas y (iv) en un mejor tratamiento de formas con prefijos y sufijos, cuya representación fonológica se hereda de las características del propio afixo.

Así mismo, para desambiguar formas homógrafas heterófonas que pertenezcan a la misma categoría morfosintáctica (y que por lo tanto, no son diferenciadas por un etiquetador morfosintáctico), pueden aplicarse gramáticas locales que tengan en cuenta el contexto semántico más próximo.

Bibliografía

- Blevins, J. 1995. The Syllable in Phonological Theory. En J. A. Goldsmith, editor, *The Handbook of Phonological Theory*. Blackwell Publishers, Cambridge, MA, páginas 206–244.
- Braga, Daniela y Luís Coelho. 2006. Letter-to-sound conversion for galician TTS systems. En *Actas de las IV Jornadas en Tecnologia del Habla*, páginas 171–176, Zaragoza.
- Braga, Daniela, Luís Coelho, y F. Resende Jr. 2006. A Rule-Based Grapheme-to-Phone Converter for TTS Systems in European Portuguese. En *Proceedings of the VI International Telecommunications Symposium (ITS'06)*, Fortaleza.
- Braga, Daniela y Xosé Ramón Freixeiro. 2007. Algoritmos de Conversão Grafema-Fone em Galego para Sistemas de Conversão Texto-Fala. En Burghard Baltrusch y Gabriel Pérez Durán, editores, *Estudos galegos de Tradución & Paratradución no século XXI*, Vigo. Xerais.
- Campillo, Francisco, Daniela Braga, Ana B. Mourín, Carmen García-Mateo, Pedro Silva, Miguel Sales Dias, y Francisco Méndez. 2010. Building High Quality Databases for Minority Languages such as Galician. En *Proceedings of the 7th conference on International Language Resources and Evaluation (LREC'10)*, Valletta.
- Carreras, Xavier, Isaac Chao, Lluís Padró, y Muntsa Padró. 2004. FreeLing: An Open-

- Source Suite of Language Analyzers. En *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC'04)*, Lisboa.
- Chomsky, Noam y Morris Halle. 1968. *The Sound Pattern of English*. Harper and Row, New York.
- García, Marcos y Pablo Gamallo. 2010. Análise Morfosintáctica para Português Europeu e Galego: Problemas, Soluções e Avaliação. *Linguamática. Revista para o Processamento Automático das Línguas Ibéricas*, 2(2):59–67.
- González González, Manuel, Eduardo Rodríguez Banga, Francisco Campillo, Francisco Méndez, Leandro Rodríguez Liñares, y Gonzalo Iglesias. 2008. Specific features of the Galician language and implications for speech technology development. *Speech Communication*, 50:874–887.
- ILG/RAG. 2005. *Normas Ortográficas e Morfolóxicas do Idioma Galego*. Real Academia Galega e Instituto da Lingua Galega, Vigo, 20 edición.
- Ladefoged, Peter. 1990. The Revised International Phonetic Alphabet. *Language*, 66(3):550–552.
- Malvar, Paulo, José Ramom Pichel, Óscar Senra, Pablo Gamallo, y Alberto García. 2010. Vencendo a escassez de recursos computacionais. Carvalho: Tradutor automático estatístico inglês-galego a partir do corpus paralelo Europarl Inglês-Português. *Linguamática. Revista para o Processamento Automático das Línguas Ibéricas*, 2(2):31–38.
- Mateus, Maria Helena Mira y Ernesto d' Andrade. 2000. *The Phonology of Portuguese*. Oxford University Press, Oxford.
- Mohanan, Karuvannur P. 1986. *The Theory of Lexical Phonology*. Reidel, Dordrecht.
- Mourín, Ana, Daniela Braga, Luís Coelho, Carmen García-Mateo, Francisco Campillo, y Miguel Dias. 2009. Homograph Disambiguation in Galician TTS Systems. En *IX Congreso Internacional da Asociación Internacional de Estudos Galegos*, A Coruña - Santiago de Compostela - Vigo.
- Oliveira, Catarina, Lurdes Castro Moutinho, y António J. S. Teixeira. 2005. On european Portuguese automatic syllabification. En *Proceedings of Interspeech 2005*, páginas 2933–2936.
- Paulo, Sérgio, Luís C. Oliveira, Carlos Mendes, Luís Figueira, Renato Cassaca, Céu Viana, y Helena Moniz. 2008. DIXI — A Generic Text-to-Speech System for European Portuguese. En *Proceedings of the 8th international conference on Computational Processing of the Portuguese Language (PROPOR'08)*, páginas 91–100.
- Regueira, Xosé Luís. 2010. *Dicionario de Pronuncia da Lingua Galega*. Real Academia Galega e Instituto da Lingua Galega, A Coruña.
- Seara, Izabel C., Fernando S. Pacheco, Rui Seara Júnior, Sandra G. Kafka, Simone Klein, y Rui Seara. 2003. Geração Automática de Variantes de Léxicos do Português Brasileiro para Sistemas de Reconhecimento de Fala. En *Actas do XX Simpósio Brasileiro de Telecomunicações*, Rio de Janeiro.
- Selkirk, E. O. 1984. On the major class features and syllable theory. En M. Aronoff y R. T. Oehrle, editores, *Language Sound Structure*. The MIT Press, Cambridge, MA, páginas 107–136.
- Siravenha, Ana C., Nelson Neto, Valquiria Macedo, y Aldebaro Klautau. 2008. Uso de Regras Fonológicas com Determinação de Vogal Tônica para Conversão Grafema-Fone em Português Brasileiro. En *Proceedings of the 7th International Information and Telecommunication Technologies Symposium (I2TS'08)*, Foz do Iguaçu.
- Teixeira, António J. S., Catarina Oliveira, y Lurdes Castro Moutinho. 2006. On the Use of Machine Learning and Syllable Information in European Portuguese Grapheme-Phone Conversion. En *Proceedings of the 7th International Conference on Computational Processing of Portuguese Language.*, páginas 212–215.
- Vigário, Marina, Fernando Martins, y Sónia Frota. 2006. A ferramenta FreP e a frequência de tipos silábicos e classes de segmentos no Português. En *Seleção de Comunicações apresentadas no XX Encontro Nacional da Associação Portuguesa de Linguística*, páginas 675–687.