

El alfabeto fonético SAMPA y el diseño de corpora fonéticamente balanceados

Como referenciar:

Gurlekian, J. A., Colantoni, L. y Torres, H. "El alfabeto fonético SAMPA y el diseño de corpora fonéticamente balanceados". *Fonoaudiológica*. Editorial: ASALFA. Tomo: 47, Numero: 3, pp 58-69, Diciembre 2001.

1. Introducción:

El presente trabajo tiene como objetivo presentar la adaptación al español del alfabeto fonético SAMPA (*Speech Assessment Methods: Phonetic Alphabet*, métodos para la evaluación del habla: alfabeto fonético) (ESPRIT, 1989) y discutir algunas de sus posibles aplicaciones.

Este alfabeto fue diseñado por un grupo de científicos europeos para codificar una notación fonética en un sistema fácil de leer en una computadora. El SAMPA asocia los símbolos del alfabeto fonético internacional con un símbolo ASCII (*American Standard Code for Information Interchange*). El resultado es un alfabeto fonético simplificado que utiliza el teclado convencional, lo cual facilita su uso por parte de quienes tengan que diseñar textos sin profundizar a priori en las variantes de la notación fonética.

El SAMPA ha sido adaptado a numerosas lenguas, incluyendo el español (Listerri y Mariño, 1993). En este trabajo, partimos de esta adaptación para incluir todos los símbolos necesarios que nos permitan transcribir todos los dialectos del español de la Argentina.

El alfabeto SAMPA revisado que aquí se propone contiene un listado completo de los fonos del español de la Argentina. Este listado puede ser utilizado en

diferentes aplicaciones, desde la simple notación fonética hasta el diseño de corpora lingüísticos. Estos son necesarios en: 1) los estudios clínicos, como las pruebas de percepción logo-audiométricas y la grabación de voces para la evaluación de los atributos de la voz y de los sonidos del habla; 2) los estudios dialectales, que requieren un texto breve que contenga los fonos posibles para permitir la comparación interdialectal; 3) la creación de bases de datos para reconocimiento, síntesis de voz y la realización de pruebas de niveles de calidad en la transmisión de voz por líneas telefónicas.

2. La adaptación del alfabeto SAMPA al español de la Argentina:

En esta sección, se discutirán las ventajas comparativas del SAMPA (ver 2.1); se propondrá la descripción fonético-articulatoria de los símbolos (ver 2.2); y se presentará la distribución dialectal de cada uno de los fonos (ver 2.3). Por último, se detallarán las diferencias entre la propuesta actual y la lista SAMPA para el español peninsular (ver 2.4).

En la Tabla I se presenta el alfabeto SAMPA para el español de la Argentina, y una lista de símbolos IPA (*International Phonetic Alphabet*) para cada uno de los

símbolos SAMPA. Además, se agrega principalmente de la propuesta de una descripción fonética sumaria tomada (Borzzone de Manrique, 1980).

Tabla I: Alfabeto SAMPA para el español de la Argentina. Las definiciones de modo y punto han sido tomadas de las siguientes referencias:

- Borzzone de Manrique (1980).
- International Phonetic Association (1999).
- Ladefoged & Maddieson (1996).

SAMPA para la Argentina					
N	IPA	Sampa	Modo, Punto, F0	Transc.	Palabra
1	i	i	Vocal, cerrada-anterior, sonora.	Bis	Bis
2	e	e	Vocal, cerrada-media-anterior, sonora.	Mes	Mes
3	a	a	Vocal, abierta-central, sonora.	Mas	Más
4	o	o	Vocal, cerrada-media-posterior-redondeada, sonora.	Tos	Tos
5	u	u	Vocal, cerrada-posterior-redondeada, sonora.	tul	Tul
6	j	j	Aproximante, palatal, sonora	laBjo	Labio
7	w	w	Aproximante, labial velar, sonora.	aGwa	Agua
8	l	l	Lateral, ápico-gingival, sonora.	loBo	Lobo
9	m	m	Nasal, bilabial, sonora	mesa	Mesa
10	ɱ	m	Nasal, labiodental, sonora.	Emfermo	Enfermo
	n	n	Nasal, ápico-gingival, sonora.	NaDa	Nada
11	ŋ	N	Nasal, velar, sonora.	ONGo	Hongo
12	ɲ	J	Nasal, dorso-palatal, sonora.	NiJo	Niño
13	β	B	Fricativa, bilabial, sonora.	TuBo	Tube
14	ð	D	Fricativa, apico-interdental, sonora.	ADa	Hada
15	ɣ	G	Fricativa, dorso-velar, sonora.	AGwa	Agua
16	b	b	Oclusiva, bilabial, sonora.	Beso	Beso
17	d	d	Oclusiva, apico-dental, sonora.	Dar	Dar
18	g	g	Oclusiva, dorso-velar, sonora.	Gula	Gula
19	r	r	Vibrante simple, ápico-gingival, sonora.	Pero	Pero
20	ɾ	R	Vibrante múltiple, ápico-gingival, sonora.	CaRo	Carro
		R	Vibrante múltiple-fricativa, apical, sonora	CaRo	Carro
21	l	L	Lateral, dorso-prepalatal, sonora.	LuBja	Lluvia
	ɮ	Z	Fricativa, dorso-prepalatal, sonora.	ZuBja	Lluvia
	ɮ̪	Z	Africada, dorso-prepalatal, sonora.	KonZuGe	Cónyuge
	ɮ̥	Z	Fricativa, dorso-prepalatal, sorda.	ZuBja	Lluvia
	ɟ	Z	Fricativa, dorso-palatal, sonora	ZuBja	Lluvia
22	h	h	Fricativa, laríngea, sorda.	ahta	Hasta
23	p	p	Oclusiva, bilabial, sorda.	pala	Pala
24	t	t	Oclusiva, ápico-dental, sorda.	tierra	Tierra
25	k	k	Oclusiva, dorso-velar, sorda.	kilo	Kilo
26	t͡ʃ	H	Africada, dorso-prepalatal, sorda.	teHo	Techo
27	s	s	Fricativa, dorso-gingival, sorda.	sala	Sala
28	f	f	Fricativa, labio-dental, sorda.	fe	Fe
29	x	x	Fricativa, dorso-velar, sorda.	xwes	Juez
30	ç	C	Fricativa, dorso-palatal, sorda.	arCentina	Argentina

2.1 Ventajas comparativas del SAMPA

El alfabeto fonético SAMPA tiene características peculiares que lo diferencian de otros alfabetos fonéticos, como el IPA, y que lo hacen más adecuado para el diseño de base de datos de textos. Las ventajas del SAMPA son las siguientes:

1. Utiliza un subconjunto de caracteres del alfabeto latino, a diferencia del IPA que incluye símbolos del alfabeto griego y una serie de marcas diacríticas. Este subconjunto de caracteres se corresponde con los caracteres ASCII y puede ser leído y reconocido en cualquier medio de procesamiento digital
2. Presenta una lista reducida de símbolos, puesto que se emplea para realizar transcripciones fonéticas amplias, que reflejen sólo los alófonos en distribución complementaria, es decir los alófonos que se pueden predecir a partir del contexto fonológico. Así un símbolo SAMPA puede equivaler a más de un símbolo IPA. Por ejemplo, en el caso de las palatales, el símbolo SAMPA 'Z' representa a cuatro símbolos IPA.
3. La utilización del SAMPA permite diseñar bases de datos de textos en las que sea necesario asegurar la presencia de todos los alófonos en distribución complementaria. Para esta tarea se

requiere la utilización de símbolos que predigan los fonemas y sus variantes, sin determinar a priori su realización concreta. El IPA carece de esta flexibilidad pues exige símbolos que designen a los sonidos de una forma no ambigua (International Phonetic Association, 1999), 3), i.e. que cada sonido se corresponda exactamente con un símbolo.

2.2 Descripción fonético-articulatoria de los fonos consonánticos del alfabeto SAMPA:

La presentación del alfabeto SAMPA motivó la revisión de las caracterizaciones fonéticas propuestas para cada uno de los fonos del español. Observamos que, si bien existían detalladas descripciones articulatorias para el español peninsular (Navarro Tomás, 1970), (Quilis, 1993), entre otros) basadas en palatogramas o radiografías, las descripciones articulatorias para el español de la Argentina eran escasas y básicamente impresionistas. (Borzzone de Manrique, 1980): 77) presenta una tabla con caracterizaciones articulatorias en la que se contrastan descripciones para el español peninsular y el español de la Argentina. Esta descripción, parece estar basada en una combinación de las conclusiones de otros autores y de observaciones realizadas por la autora. Decidimos, no obstante, utilizarla por ser la única disponible para nuestro español.

Señalaremos aquí qué aspectos deben ser revisados, teniendo en cuenta las nuevas tendencias en fonética articulatoria. Así, nuestro objetivo es presentar los problemas que se plantean en la descripción de ciertos fonos, especialmente en los fonos consonánticos, y abrir el camino para nuevos estudios.

En las caracterizaciones articulatorias del español se observan, en general, dos problemas principales: una imprecisión en las descripciones, acompañada por una falta de discusión acerca de la naturaleza de las caracterizaciones articulatorias; y una separación entre las caracterizaciones fonéticas y la fonología del español, i.e. las descripciones no toman normalmente en cuenta el papel que los segmentos juegan en la fonología del español. Esta separación, hace que las descripciones fonéticas no expliquen por qué algunos sonidos pueden tener articulaciones menos precisas, mientras que a la fonología le hace difícil dar cuenta de procesos de alternancias ampliamente difundidos - como la espirantización - pero no debidamente explicados, dada la imprecisión de las descripciones articulatorias en el español. (Colantoni, 2001: 73-76).

En cuanto a la primera limitación, hay que destacar que las descripciones existentes no presentan una caracterización precisa de los articuladores activos y pasivos. Los autores españoles no coinciden, por ejemplo, en la caracterización de las partes del principal articulador activo, la lengua. Así, para Navarro Tomás (Navarro Tomás, 1970): ¶11) y (Quilis, 1993), la lengua se divide en ápice, predorso, mediodorso, postdorso y raíz,

mientras que para Borzone de Manrique (Borzone de Manrique, 1980): 60) se divide en punta, dorso y raíz. Por otra parte, Ladefoged y Maddieson (Ladefoged, 1996) distinguen en ella cuatro zonas principales: ápice, lámina¹, cuerpo y raíz. Los autores observan que, aunque hay acuerdos acerca de lo que constituye el centro de cada una de estas áreas, no siempre es sencillo trazar los límites entre cada una de ellas. Otro tanto ocurre con los articuladores pasivos, como las distintas secciones del paladar duro.

Otro aspecto de la descripción articulatoria a la que no se le ha prestado la debida atención es el papel de la configuración que adopta la lengua en la articulación, (Stone et al, 1992; Stone y Vatikiotis-Bateson, 1995) especialmente relevante en los sonidos alveolares y palatales. Stone et al. (1992, 268) muestran que en [ʃ], [l] y [s] la parte anterior de la lengua adopta una configuración diferente a la de la sección posterior.

A continuación, revisaremos la caracterización de las consonantes. En esta revisión, que se organizará de acuerdo con los puntos de articulación (labial, dental, alveolar, palatal y velar), señalaremos: las principales divergencias, las revisiones que deben efectuarse y, por último, los avances que se hayan realizado.

Las consonantes labiales ([p], [b], [B], [m], [f]) y las velares ([k], [g], [G], [x], [N]) son las menos problemáticas en la caracterización del punto de articulación,

¹ Zona contigua al ápice en la parte anterior, que traducimos del término inglés *blade*.

puesto que éste se localiza en algún extremo del continuo del tracto vocal. En particular en el caso de las labiales, los articuladores activos (labios) y pasivos (dientes, para el caso de las labiodentales) son directamente observables.

Ha habido, en cambio, discrepancias en la caracterización del modo de articulación de la labial sonora [B], y la velar sonora [G]. La mayoría de los autores (Navarro Tomás, 1970); (Borzone de Manrique, 1980); (Quilis, 1993) las considera fricativas, al igual que a la dental sonora [D]. Se ha demostrado, sin embargo, que no se trata de sonidos fricativos, sino ‘aproximantes’ ya que no se ha observado ruido de fricación (Santagada y Gurlekian, 1989).

El mayor problema en la caracterización de las consonantes dentales ([t], [d], [D]) y en el de las alveolares ([n], [l], [s], [r], [R]) es el papel del articulador activo. (Navarro Tomás, 1970): ¶98 sostiene que éste es la “punta de la lengua”, sin precisar (ibid ¶ 11) el límite entre el ápice y el predorso. Algo similar se observa en la caracterización de (Quilis, 1993): 197), y (Borzone de Manrique, 1980) quienes consideran al ápice de la lengua como el articulador activo, sin indicar si esta zona abarca también la lámina (ibid. 64).

Por otro lado, (Ladefoged, 1996): 20-21), basándose en la comparación de numerosas lenguas, concluyen que la lámina de la lengua es el articulador activo en las dentales, mientras que el ápice lo es en el caso de las alveolares. Es importante, entonces, revisar las descripciones existentes, partiendo de una descripción más precisa de los

articuladores basada en estudios que empleen la tecnología adecuada.

Las consonantes de la región prepalatal o palatal ([L], [Z], [H], [C] [J]) son las más conflictivas, sin lugar a dudas. Las divergencias en la caracterización se relacionan tanto con el modo como con el punto de articulación. Los autores no concuerdan sobre si el articulador activo es el ápice, el predorso o el dorso de la lengua. Por ejemplo, en la caracterización de la nasal palatal [J], (Navarro Tomás, 1970)(1970) señala que interviene la punta y el dorso; (Quilis, 1993)) indica que el predorso es el articulador móvil, y (Borzone de Manrique, 1980) le atribuye al dorso de la lengua el papel activo en la articulación. Otro tanto ocurre con la determinación de la zona del articulador fijo que interviene. En el mismo caso de la palatal nasal, Navarro Tomás (ibid.) señala que la zona de contacto se extiende desde los alvéolos hasta el post-paladar, mientras que Quilis (ibid.) y Borzone de Manrique (ibid.) sostienen que el contacto se da en la zona prepalatal.

Estas divergencias, como ya mencionamos, se deben a que (i) no hay coincidencia en la denominación de las partes de la cavidad oral; (ii) no se tiene en cuenta la posición que adopta el cuerpo de la lengua en la articulación; (iii) no se considera que existe un continuo difícil de delimitar que va desde la región postalveolar a la postpalatal; (iv) no se contemplan las posibilidades que deja abiertas el sistema fonológico del español. De hecho, el español, en la mayoría de sus dialectos, tiene un solo fonema oral fricativo palatal ([Z]) que sólo se opone a una africada palatal. El

área palatal es más extensa comparada con la dento-alveolar y, sin embargo, existen menos fonemas que en ella. Así, la posibilidad para la variación fonética es más amplia en el área palatal que en otras (Guitarte, 1955: 261-283), (Colantoni, 2001). Por estos motivos, las consonantes palatales son las que presentan mayor variación en puntos de articulación y en modos de articulación en los dialectos geográficos y sociales.

2.3. Variaciones dialectales en el español de la Argentina:

Ante todo, es importante destacar que en el español de la Argentina existen dos inventarios fonémicos, puesto que en algunas zonas se distinguen, en la pronunciación, los grafemas ‘ll’ e ‘y’. Estas zonas abarcan parte de la región mesopotámica (Corrientes, Misiones, este del Chaco y de Formosa y noreste de Santa Fe), una parte de la región cordillerana (noroeste de La Rioja, oeste de Catamarca y norte de San Juan), parte de la provincia de Jujuy y la provincia de Santiago del Estero (Vidal de Battini, 1964).

En todas estas zonas, excepto en Santiago del Estero (Vidal de Battini, 1964), el grafema ‘ll’ se realiza como una palatal lateral [L], mientras que el grafema ‘y’ se realiza ya como una africada palatal sonora [d̪j] en la región mesopotámica, ya como una semiconsonante [j] o una fricativa palatal [j̥], en la zona andina². Es importante, sin

embargo, mencionar que el estado de esta oposición no es estable. Se ha observado (Colantoni, 2001) que se mantiene en un alto porcentaje en el centro y norte de Corrientes, pero muestra signos de desaparición en el sur, mientras que ha desaparecido prácticamente por completo en el norte de San Juan, donde la distinción sólo se mantiene en ítems léxicos aislados.

En la provincia de Santiago del Estero, la oposición tiene una realización fonética diferente. El grafema ‘ll’ se realiza como una palato-alveolar fricativa sonora [ʒ], mientras que el grafema ‘y’ se realiza como una palatal fricativa sonora [j̥].

Las restantes variaciones que se observan en el español de la Argentina tienen que ver con la realización de ciertos alófonos de las consonantes palatales, de las vibrantes y de la alveolar fricativa sorda [s].

Las consonantes palatales [Z] son las que presentan mayor cantidad de variantes, entre las que se han observado las siguientes:

- a. [ʒ]: (Vidal de Battini, 1964) la registra en la ciudad y en la provincia de Buenos Aires, la totalidad de la Patagonia, sur de Córdoba, Santa Fe, Entre Ríos y Tucumán. (Alvar, 1995a: 125-145) nota su presencia en Salta; (Coronel, 1995: 44-52)) observa que se está extendiendo en Catamarca y (Colantoni, 2001) señala su extensión en Córdoba, San Luis, Entre Ríos y

² En todos los casos, se utilizará el símbolo SAMPA [Z] para transcribir el grafema ‘y’, aunque su descripción fonética variará de acuerdo

con la realización más frecuente que se observe en cada dialecto.

- algunas localidades de la provincia de Corrientes.
- b. [ʃ]: Ha sido observada en la ciudad de Buenos Aires (Zamora Vicente, 1949: 5-22); (Guitarte, 1955: 261-283); (Martín, 1980), (Wolf, 1979: 115-145; Wolf, 1984: 175-196), en la provincia de Buenos Aires (Fontanella de Weinberg, 1979), en el sur de Córdoba, San Luis y Entre Ríos (Colantoni, 2001).
 - c. [dʒ]: Presente en la ciudad de Buenos Aires (Wolf, 1979: 115-145; Wolf, 1984: 175-196) y en baja frecuencia en algunas localidades de la provincia de Córdoba y Entre Ríos (Colantoni, 2001).
 - d. [j]: Se registra en gran parte del territorio argentino, especialmente en la zona central y el noroeste (Vidal de Battini, 1964; Colantoni, 2001).
 - e. [d̪]: Su extensión se reduce a la zona de la Mesopotamia (Vidal de Battini, 1964), especialmente a la zona del centro y norte de la provincia de Corrientes (Colantoni, 2001).

La vibrante múltiple [R], presenta dos realizaciones principales:

- a. [r]: Es la variante que se registra en la provincia de Buenos Aires, la Patagonia, sur de Córdoba, Santa Fe (Vidal de Battini, 1964). Alvar (Alvar, 1995b: 125-145) observa su extensión en la provincia de Salta, mientras que Colantoni (Colantoni, 2001) señala una alta frecuencia de uso en la mayor parte de la provincia de Córdoba, en San Luis, Entre Ríos y sur de Corrientes.

- b. [r̄]: Esta variante se observa en el resto de las provincias argentinas. La realización de este sonido puede ser sonoro, como en el caso de la región mesopotámica, o presentar ensordecimientos, como en la provincia de San Juan (Colantoni, 2001).

Por último, la /s/, el fonema más estudiado, presenta tres variantes principales, de acuerdo con la región, el grupo social y el contexto fonológico³.

- a. [s]: Vidal de Battini (1964) observa una tendencia hacia el mantenimiento de [s] en la provincia de Buenos Aires y en la Patagonia. Terrell (1979) la considera la variante más frecuente en la ciudad de Buenos Aires. Fontanella de Weinberg (1967) observa un mantenimiento ante vocal y ante dentales en variación libre con los alófonos que se detallan en (b) y (c) en la ciudad de Bahía Blanca. Finalmente, en Santiago del Estero se mantiene una [s] tensa (Vidal de Battini, 1964).
- b. [h]: Kovacci y Martín (1970) y Borzone de Manrique (1980) observan la presencia de aspiración en Buenos Aires, fundamentalmente ante consonantes sordas. Terrell (1979) señala que, en el habla culta de Buenos Aires, es la variante menos frecuente. Donni de Mirande (1986) indica que, en el español de Rosario, la aspiración se da en posición interior de palabra. Por último, Kovacci (1987) documenta la presencia de la

³ La /s/ presenta variaciones más significativas en posición implosiva silábica, tanto en posición interna como en posición final de palabra. A estas variaciones nos referimos aquí.

aspiración en la provincia de Corrientes.

- c. [0]: Según Vidal de Battini (1964), la región guaranítica se caracteriza por una clara tendencia a la elisión en posición final de sílaba y final de palabra. Donni de Mirande (1986) observa esa misma tendencia en la ciudad de Rosario, y Terrell (1979) indica que, en la ciudad de Buenos Aires, la elisión supera a la aspiración.

2.4. Adaptación del alfabeto fonético SAMPA para el español de la Argentina:

Dadas las particularidades descriptas anteriormente para el español de la Argentina, fue necesario adaptar la versión SAMPA del español peninsular realizando modificaciones para incluir todos los símbolos necesarios que nos permitieran transcribir todos los dialectos del español de la Argentina. Además, se efectuaron simplificaciones de símbolos compuestos por dos caracteres a uno por razones de consistencia interna.

2.4.1 Modificaciones:

Los agregados se relacionan con las variantes fonéticas específicas del español de la Argentina. Se agregaron tres símbolos que aparecen descriptos a continuación:

- a. [Z]: se emplea para transcribir la palatal fricativa sonora [ʝ], la palatal africada sonora [d͡ʝ], la palato-alveolar

fricativa sonora [ʒ], la palato-alveolar fricativa sorda [ʃ], y la palato-alveolar africada sonora [d͡ʒ]. La primera variante se registra, por ejemplo, en el español de Córdoba, Cuyo, el Noroeste y parte del Noreste de la Argentina. La segunda variante se observa en la zona de Corrientes y Misiones, en donde es un fonema que contrasta con la palatal lateral. Las tres últimas variantes son propias del español de Buenos Aires, aunque se registran también en el sur de Córdoba, Santa Fe, Entre Ríos y la totalidad de la Patagonia.

b. [h]: representa a la fricativa glotal sorda, alófono de [s] en posición implosiva silábica. En la Capital Federal (Borzzone de Manrique, 1980), esta variante se registra sólo ante consonante, mientras que en otras zonas de la Argentina, se observa tanto en posición interna como posición final de palabra (Vidal de Battini, 1964).

c. [C]: se emplea para transcribir la palatal fricativa sorda, que es un alófono de la velar fricativa sorda cuando ésta está seguida por una vocal anterior. Esta variante ha sido descripta para el español de la ciudad de Buenos Aires (Borzzone de Manrique, 1980).

Por último, es importante destacar que se conservó el símbolo [L] para transcribir la palatal lateral⁴ que se registra en los dialectos del norte de la Mesopotamia, mientras que se eliminaron

⁴ En la transcripción del dialecto hablado en Santiago del Estero, estos símbolos tienen un valor fonético diferente. El grafema 'll' se transcribe con el símbolo [L], aunque su realización fonética sea una palato-alveolar fricativa sonora. Por otro lado, se emplea el símbolo [Z] para transcribir el grafema 'y', realizado como una fricativa palatal.

símbolos como [θ]⁵, que no es de uso general en nuestros dialectos.

2.3.2 Simplificaciones:

Las simplificaciones consistieron en la reducción a un solo carácter de aquellos símbolos representados por dos caracteres en la lista SAMPA original. Así, los símbolos que aparecen a la izquierda fueron reemplazados por los de la derecha:

[tS] → [H]

[rr] → [R]

3 El transcriptor ortográfico

El listado completo del SAMPA para el español de la Argentina se incorporó a un transcriptor ortográfico-fonético automático, desarrollado en el Laboratorio de Investigaciones Sensoriales. El sistema tiene dos componentes: un transcriptor ortográfico-fonético (Figura 1) y un módulo que se aplica luego de realizada la transcripción, para calcular la frecuencia de los símbolos SAMPA en el texto de origen (Figura 2).

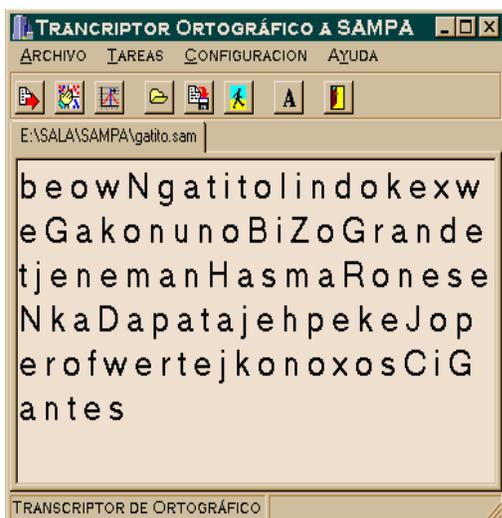


Figura 1. Visualización de la pantalla de salida del transcriptor automático de texto a SAMPA

⁵ Se ha observado como variante de la /s/ en algunas regiones de la Argentina. (Vidal de Battini, 1964; Kovacci, 1987 y Alvar, 1995).



Figura 2. Visualización de la pantalla de salida con el número de ocurrencias de cada símbolo SAMPA para el texto 5 de la sección 5.

El primer componente permite transcribir automáticamente cualquier texto de entrada en un texto con símbolos SAMPA. Esta transcripción contiene las variantes fonéticas que se pueden predecir a partir del contexto, es decir, sólo se transcriben aquellas variantes que se derivan a partir de la posición del grafema en la cadena del texto, sobre la base de los conocimientos adquiridos acerca de la distribución de las consonantes y vocales del español. Estos conocimientos se emplearon para crear un conjunto de reglas que transforman un grafema en un símbolo SAMPA (Ver Tabla II).

Tabla II. Relación entre grafemas y símbolos SAMPA de acuerdo con los diferentes contextos. (- : grafema, #: pausa, C: consonante, V: vocal, -ac grafema vocálico acentuado).

Grafema	Contexto	SAMPA
i	# -C, C-C, -ac V V-, -V	i j
e	Todos	e

a	Todos	a
o	Todos	o
u	# -C, C-C, -ac V V-, -V	u w
b	# -, C _{nasal} - Restantes	b B
c	-V _{e, i} -V _{a, o, u}	s k
ch	Todos	H
d	# -, C _{nasal} -, C _l - Restantes	d D
f	Todos	f
g	# -C, C _{nasal} - -V _{e, i} Restantes	g C G
h	Todos	-
j	-V _{o, a, u} -V _{e, i}	x C
k	Todos	k
l	Todos	l
ll	Todos	Z
m	Todos	m
n	-C _f -C _{velar} Restantes	m N n
ñ	Todos	J
p	Todos	p
q	Todos	k
r	# -, C _{n, r, s} Restantes	R r
rr	Todos	R
s	-C Restantes	h s
t	Todos	t
v	# -, C _{nasal} - Restantes	b B
w	Todos	w
x	Todos	k s ^l
y	-V	Z
z	-C Restantes	h s

La transcripción se realiza oración por oración. Las reglas para la transcripción fonética se aplican a cada palabra y, algunas de ellas, se aplican entre palabra y palabra. Estas reglas, que dan lugar a las sílabas fusionadas (Guirao y García Jurado, 1993), eliminan el último sonido de una palabra (generalmente una vocal) y lo unen al inicio de la palabra siguiente.

Una vez realizada la transcripción, se utiliza el segundo componente del programa, que permite calcular las frecuencias de ocurrencia de los fonos obtenidos. Este componente permite verificar, además, si todos los símbolos de la lista SAMPA se encuentran presentes en el texto transcripto.

5 Aplicación del diseño de corpora fonéticamente balanceados en Foniatría.

El alfabeto SAMPA, implementado en un software de transcripción automática, es una herramienta importante para la creación de corpora fonéticamente balanceados que contengan todos los fonos de una determinada lengua o dialecto.

Los corpora encuentran su aplicación en pruebas de percepción y producción que requieren material rico fonéticamente y estandarizado para permitir la comparación y normalización de los resultados obtenidos en un mismo sujeto y entre sujetos.

El listado SAMPA implementado en el transcriptor se utilizó para diseñar y evaluar cinco textos cortos que no superen los quince segundos de duración.

Estos textos, que contienen todos los alófonos del español de Buenos Aires, están destinados a ser leídos por hablantes de distinta edad. El texto de entrada es transformado a símbolos SAMPA y analizado estadísticamente para mostrar la frecuencia de ocurrencia de cada uno de los símbolos. A continuación, se transcribe ortográfica y fonéticamente

cada uno de los textos producidos para la evaluación.

1. Digo rápido algunas palabras fáciles en español, o castellano, justamente la lengua que se habla en la Argentina. Bueno, muchas gracias.

diGoRápidoalGunahpalaBrasfásilesenehpaJolokahteZano
xuhmentelaleNgwakeaBlaenlaarCentina
bwenomuHasGrasjas

2. Bueno, ya no sé qué pasa con mi pobre coche viejo. Tengo fuego en el caño de escape, y cuando arranca, se prende la luz de giro.

bwenoZanosekepasaKonmipoBrekoHeBjexoteNgofweGoe
nelkaJoDeehkapejkwandoaRaNkaseprendelalusDeCiro

3. Te llamo para cambiar el horario de la reunión de mañana. ¿Podemos encontrarnos a las ocho? Tengo que ir a la escuela de Gerardo porque juega la final de futbol.

teZamoparakambjarelorarjoDelaRewnjondeMaJanapode
mosenKontrarnosalasoHoteNgokejralaehkwelaDeCerard
oporkexweGalafinaldefutBol

4. ¡Betty! ¡Qué inmensa alegría escucharte! Cuando vengas para fin de año, quiero llevarte a recorrer toda la Argentina.

bettikejnmensaaleGriaehkuHartekwandoBeNgaHparafinde
aJokjeroZeBartearekoRertoDalaarCentina

5. Veo un gatito lindo que juega con un ovillo grande. Tiene manchas marrones en cada pata y es pequeño pero fuerte y con ojos gigantes.

beowNgatitolindokexweGakonunoBiZoGrandetjenemanHa
smaRoneseNkaDapatajehpekeJoperofwertekonoxosCiGan
tes

6 Conclusiones:

El presente artículo expone el origen y las ventajas del alfabeto SAMPA para el diseño de corpora fonéticamente balanceados. Su principal característica es

que permite diseñar bases de datos de textos en las que sea necesario asegurar la presencia de todos los alófonos en los diferentes contextos fonológicos.

La adaptación del SAMPA para los distintos dialectos del español de la

Argentina sugirió, además, la necesidad de revisar las caracterizaciones articulatorias de los fonos consonánticos mediante técnicas dinámicas específicas.

Resumen

El presente artículo se plantea tres objetivos: mostrar las ventajas del alfabeto SAMPA, exponer sus posibles aplicaciones y, por último, sugerir la necesidad de revisar las caracterizaciones de las consonantes.

El empleo del alfabeto SAMPA presenta una serie de ventajas para el diseño de corpora fonéticamente balanceados. Los símbolos SAMPA son fáciles de aprender, dada su semejanza con el alfabeto latino, y sencillos de implementar con un teclado convencional de computadora. La utilización del SAMPA permite diseñar bases de datos de textos en las que sea necesario asegurar la presencia de todos los alófonos en los diferentes contextos fonológicos, puesto que para esta tarea se requieren símbolos que predigan los fonemas y sus variantes, sin determinar a priori su realización concreta.

Se muestra aquí la adaptación del alfabeto SAMPA para los distintos dialectos del español de la Argentina y su incorporación a un transcriptor ortográfico-fonético automático. Se indica su aplicación para el diseño de textos para pruebas foniátricas de producción y percepción.

Por último, se presenta la distribución dialectal de cada símbolo y se sugiere la necesidad de revisar las caracterizaciones

articulatorias de los fonos consonánticos debido a las imprecisiones descriptivas.

Abstract

The article has three goals: to present the advantages of the SAMPA phonetic alphabet, to show its applications, and, finally, to suggest the need for reviewing the characterization of Spanish consonants.

The SAMPA alphabet has several advantages in the design of phonetically balanced databases. SAMPA symbols are easy to learn, given the similarity with the Latin alphabet, and practical to use in a standard keyboard. The SAMPA alphabet allows the design of text databases, in which the presence of every contextual allophone of a given dialect is required, since SAMPA symbols predict the presence of all the allophones, without *a priori* determining their concrete realization.

It is shown here the application of the SAMPA in the design of text database for speech production and perception tests.

The article also deals with the adaptation of the SAMPA alphabet to every Argentine Spanish dialect. A dialectal distribution of each SAMPA symbol is presented, and it is argued that the articulatory characterization of Spanish consonants needs to be reviewed.

Bibliografía

Alvar, Manuel. Muestras de polimorfismo en el español de la Argentina. La lengua española y su expansión en la época del Tratado de Tordesillas. 125-145. 1995. Sociedad V Centenario-Tratado de Tordesillas.

- Borzone de Manrique, A. M., 1980. Manual de fonética acústica. Buenos Aires: Hachette.
- Colantoni, Laura. Mergers, chain shifts, and dissimilatory processes: palatals and rhotics in Argentine Spanish. 2001. University of Minnesota. Tesis doctoral.
- Coronel, A., 1995. Variación palatal en Catamarca, Argentina. *Missouri Philological Association* 20, 44-52.
- Donni de Mirande, N., 1986. Problemas y estado actual de la investigación del español de la Argentina hasta 1984. *Anuario de Letras* 24, 179-236.
- ESPRIT. 1989. <http://www.phon.ucl.ac.uk/home/sampa/home.htm>
- Fontanella de Weinberg, Ma. B., 1967. La -s postapical en la región bonaerense. *Thesaurus* 22, 394-400.
- Fontanella de Weinberg, Ma. B., 1979. Dinámica social de un cambio lingüístico. México: UNAM.
- Guirao, Miguelina y García Jurado, María Amalia. Estudio estadístico del español. 1993. Buenos Aires, CONICET.
- International Phonetic Association, 1999. *Handbook of the International Phonetic Association*. Cambridge: Cambridge University Press.
- Kovacci, Ofelia y Martín, Herminia. La -s en el español de Buenos Aires. *Proceedings and papers of the XIIth International Congress of AULLA*. 1970. Victoria, Clayton.
- Kovacci, O., 1987. Documentos del PREDAL Argentina 2. Atlas lingüístico-antropológico de la República Argentina. Buenos Aires: Ministerio de Educación y Justicia.
- Kovacci, O., 1987. Algunos rasgos fonéticos y gramaticales observados en la provincia de Corrientes. En: Kovacci, O. (ed.), *Documentos del PREDAL Argentina 2. Atlas lingüístico-antropológico de la República Argentina*, 14-20. Buenos Aires: Ministerio de Educación y Justicia.
- Ladefoged, P. y Maddieson, I., 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.
- Llisterri, J, y Mariño, José. Spanish Adaptation of SAMPA and automatic phonetic transcription. Informe SAM-A/UPC/001/VI. Febrero 1993.
- Martín, E., 1980. La teoría fonológica y el modelo de estructura compleja. Madrid: Gredos.
- Navarro Tomás, T., 1970. *Manual de pronunciación española*. Madrid: CSIC.
- Quilis, A., 1993. *Tratado de fonología y fonética españolas*. Madrid: Gredos.
- Santagada, M. A., Gurlekian, J. A., 1989. Spanish Voiced Stops in VCV contexts: Are they fricative variants or approximants?, *Revue de Phonetique Appliquee*, N° 90-91-92, pp. 90-100
- Stone, M., Faber, A., Raphael, L., Shawker, T., 1992. Cross-sectional tongue shape and linguopalatal contact patterns in [s], [ʃ], and [l]. *Journal of Phonetics* 20, 253-270.
- Stone, M., Vatikiotis-Beteson, E., 1995. Trade-offs in tongue, jaw, and palate contributions to speech production. *Journal of Phonetics* 23, 81-100.
- Terrell, T., 1979. La aspiración y elisión de /s/ en el español porteño. *Anuario de Letras* 16, 41-66.
- Vidal de Battini, B., 1955. Extensión de la RR múltiple en la Argentina. *Filología* 3, 181-184.
- Vidal de Battini, B., 1964. *El español de la Argentina*. Buenos Aires: Consejo Nacional de Educación.
- Wolf, C. y Jiménez, E., 1979. El ensordecimiento del yeísmo porteño. En: Barrenechea, A. Ma. (ed.), *Estudios lingüísticos y dialectológicos*, 115-145. Buenos Aires: Hachette.
- Wolf, C., 1984. Tiempo real y tiempo aparente en el estudio de una variación lingüística: ensordecimiento y sonorización del yeísmo porteño. En: Schwarz, L., Lerner, I. (eds.), *Homenaje a Ana Ma. Barrenechea*, 175-196. Madrid: Castalia.
- Zamora Vicente, 1949. Rehilamiento porteño. *Filología* 1, 5-22.

Ing. Jorge A. Gurlekian
Investigador del CONICET

Dra. Laura Colantoni
Investigadora Visitante

Ing. Humberto Torres
Becario del FONCyT

Laboratorio de Investigaciones
Sensoriales (LIS).

(54-11)-5950-9024
jag@fmed.uba.ar