

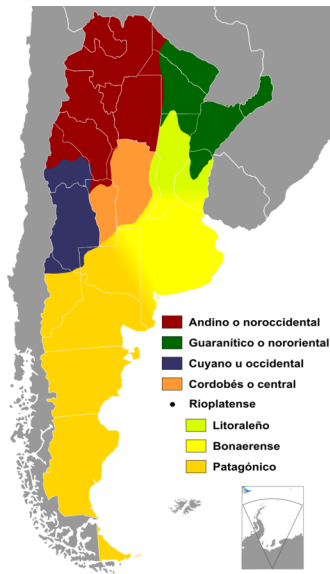
# Recolección online de grabaciones para el estudio de las variantes argentinas del español

Fernando Bugni

Departamento de Computación - Facultad de Ciencias Exactas -  
Universidad de Buenos Aires

2014

# Variantes del español en Argentina



# Diferencias entre Córdoba y Buenos Aires

- **Regla 1: Los hablantes de Córdoba estiran la sílaba anterior a la acentuada mientras los de Buenos Aires no lo hacen**

*‘Especta**cu**lar’*

Sílaba acentuada en *‘-lar’*

La sílaba anterior *‘-cu-’* se alarga para hablantes de Córdoba

# Diferencias entre Córdoba y Buenos Aires

- **Regla 2: Los hablantes de Córdoba aspiran y elisionan la /s/ al finalizar una palabra. Esto no sucede en Buenos Aires**

*'Pájaros'*

/s/ se acorta su duración en el hablante de Córdoba

# Diferencias entre Córdoba y Buenos Aires

- **Regla 3: Para hablantes de Córdoba, la /s/ antes de la /c/ o /t/ suenan más suaves que para hablantes de Buenos Aires**

*‘Mosca’*

/s/ suena más suave para Córdoba que para Buenos Aires

# Diferencias entre Córdoba y Buenos Aires

- **Regla 4: La 'c' antes de la 't' se pronuncia con menor frecuencia para hablantes de Córdoba que para hablantes de Buenos Aires**

*'Doctor'*

No debe sonar el fonema /c/

# Diferencias entre Córdoba y Buenos Aires

- **Regla 5: Para hablantes cordobeces la 'y' y 'll' se pasa a 'i'. No sucede esto para Buenos Aires**

*'lluvia'*

Palabras con el fonema /y/ o /ll/ se pronuncian /j/

# Diferencias entre Córdoba y Buenos Aires

- **Regla 6: En hablantes cordobeces la /r/ no vibra mientras que en Buenos Aires pasa lo contrario**

*‘Espárrago’*

Para Córdoba /r/ debe ser suave en comparación de Buenos Aires



# Diseño del experimento

- **Frases Comúnes:** habla espontánea
- **Frases Amper:** reconocer palabra acentuada

# Diseño del experimento

## Frases Comunes

Pronunciar frases popularmente conocidas

- Objetivo: pronunciación espontánea
- Reglas a cubrir: 2 a 6

# Diseño del experimento

**‘En la pelea se conoce al soldado,  
sólo en la **victoria** se conoce al **caballero**’**

- **‘victoria’** cubre la regla 4 que nos propone medir la duración de la /c/ antes de la /t/.
- **‘caballero’** para la regla 5: el fonema /ll/ se pasa a /i/

# Diseño del experimento

## Frases Amper

Pronunciar frases con una estructura fija variando acentuaciones

- Objetivo: cubrir acentuaciones
- Regla a cubrir: 1

*Sujeto + "salió" + Adjetivo*

- Sujeto: "El canapé", "El repollo", "El espárrago".
- Adjetivo: "espectacular", "delicioso", "riquísimo".

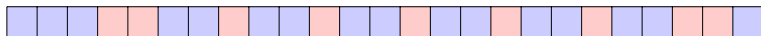
# Diseño del experimento

**“El canapé salió delicioso”**

- Canapé: palabra aguda
- Delicioso: palabra grave

# Diseño del experimento

Trazas: combinación de frases



 Frases comunes

 Frases Amper

Intercalado: 1 ó 3 Frases comunes cada una Amper

# Diseño del experimento

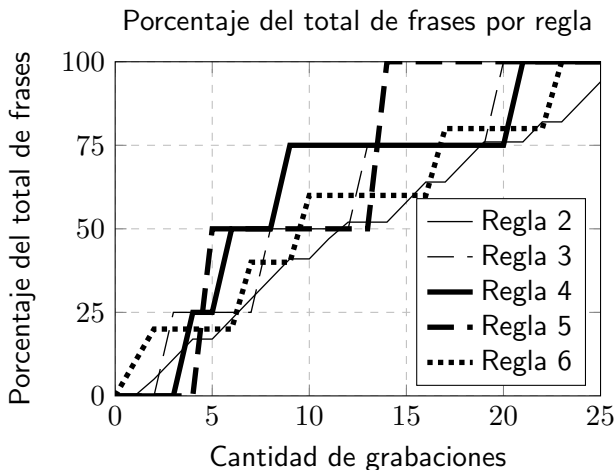


Figura : Porcentaje del total de frases grabadas por cada regla

# Sistema de grabación online

**¡Bienvenido/a!**

Este proyecto consiste en **grabar una serie de frases a través de tu computadora**, para luego poder estudiar las características del habla de cada región (por ejemplo, la tonada o los sonidos empleados).

Requisitos para poder participar:

1. Tener una **buena conexión a Internet**; preferentemente, no wireless.
2. Tener un **buen micrófono**; preferentemente, no usar el micrófono incluido en una laptop.
3. Estar en un **ambiente silencioso**.

Si cumples estos requisitos, por favor completá los siguientes datos para comenzar:

Sexo:

Lugar donde te criaste:

Lugar donde vivís actualmente:

Mes de nacimiento:  01-1990

**¡Empezar!**

Figura : Encuesta inicial del sistema



# Sistema de grabación online



Figura : Explicación del experimento

# Sistema de grabación online

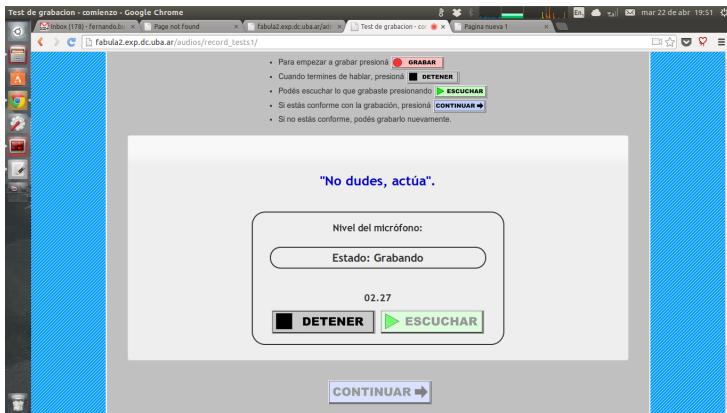


Figura : Grabando

# Sistema de grabación online

## Audio 6

**Id:** 6

**Speaker:** 2

**Word:** No está mal ser bella; lo que está mal es la obligación de serlo

**Attempt:** 1

**Filename:**



download: [bsas\\_u2\\_t32\\_a1](#)

**Labels:**

- ☐ Conservar
- ☒ Sonido saturado
- ☐ Mucho ruido de fondo
- ☐ Problema en el habla

Submit

Figura : Administrador

# Datos obtenidos

	<b>Bs.As.</b>	<b>Cba.</b>	<b>Total</b>
<b>Conservar</b>	220	90	310
<b>Problemas en el habla</b>	33	15	48
<b>Mucho ruido de fondo</b>	2	12	14
<b>Sonido saturado</b>	2	0	2

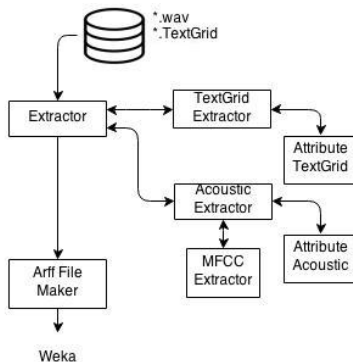
Tabla : Evaluación manual de las grabaciones

	<b>Bs.As.</b>	<b>Cba.</b>	<b>Total</b>
<b>Todos los intentos</b>	220	90	310
<b>Último intento</b>	<b>181</b>	<b>79</b>	<b>260</b>

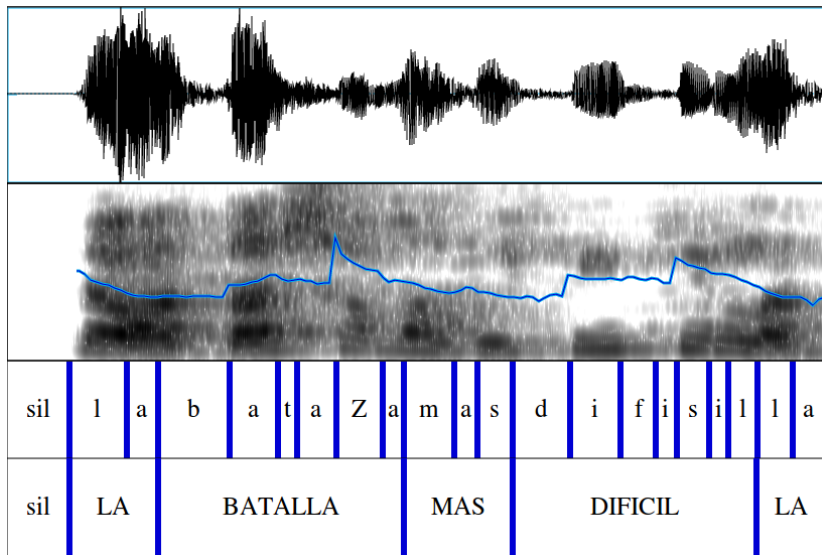
Tabla : Cantidad de audios repetidos

# Extracción de información

- ProsodyLab
- Python 2.7, SciPy y Numpy
- Pymatlab



# Extracción de información



# Extracción de información

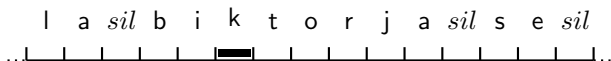
## Atributos fonéticos

- Duración de 'kt'
- Duración de 'sc'
- Duración de 'll'
- Duración de 'rr'
- Duración de 's' final
- Duración de cada fonema
- Duración de cada vocal
- Duración de cada consonante

# Extracción de información

Atributos fonético: cálculo duración de 'kt'

**“en la pelea se konose al soldaDo solo en la  
biktorja se konose al kaBaZero”**



$$\frac{X - \mu}{\sigma}$$

- $X$  es el valor a normalizar (por ej.: la duración de un fonema dado).
- $\mu$  es el promedio de duración de la unidad utilizada en la grabación.
- $\sigma$  es el desvío estándar de la unidad utilizada en la grabación.



# Extracción de información

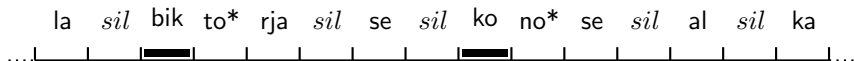
## Atributos silábicos

- Duración de la sílaba acentuada
- Duración de la sílaba anterior a la acentuada

# Extracción de información

Atributos silábico: cálculo duración de 'kt'

**“en la pelea se konose al soldaDo solo en la  
biktorja se konose al kaBaZero”**



$$\frac{X - \mu}{\sigma}$$

- $X$  es el valor a normalizar (por ej.: la duración de un fonema dado).
- $\mu$  es el promedio de duración de la unidad utilizada en la grabación.
- $\sigma$  es el desvío estándar de la unidad utilizada en la grabación.

## Clasificadores

- Zero rules
- RIPPER
- C4.5
- Support vectors machines
- Naive Bayes

## Cross-validations

- Grupos de hablantes
- Dejando un hablante fuera promediando los atributos
- Dejando un hablante fuera promediando los atributos desconocidos

# Grupo de hablantes

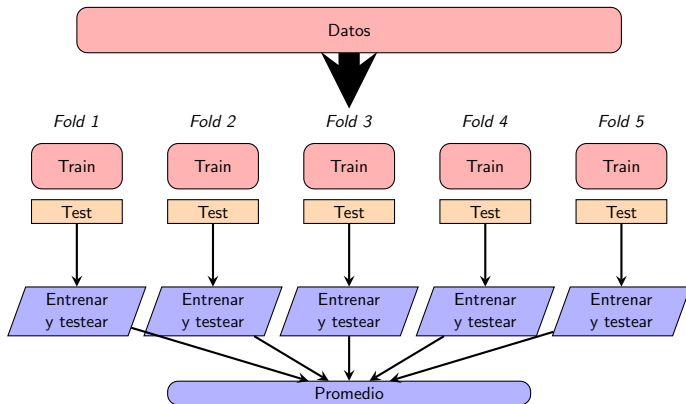


Figura : Esquema de test 5-folds

# Grupo de hablantes

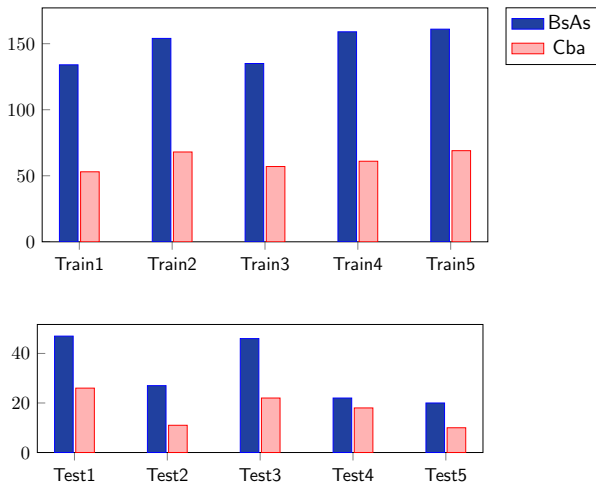


Figura : Cantidad de instancias de Buenos Aires y Córdoba según cada grupo de Train y Tests

# Grupo de hablantes

	<b>ZeroR</b>	<b>JRip</b>	<b>J48</b>	<b>Function SMO</b>	<b>NaiveBayes</b>
<b>Fold 1</b>	64	61	64	73	63
<b>Fold 2</b>	71	68	71	76	71
<b>Fold 3</b>	67	54	45	75	67
<b>Fold 4</b>	55	52	55	67	80
<b>Fold 5</b>	66	70	66	70	70
<b>Promedio</b>	64	61	60	72	70

**Tabla :** Clasificación correcta en porcentaje

	<b>Student Test</b>	<b>Wilcoxon Test</b>
<b>ZeroR y JRip</b>	0.8438	0.87
<b>ZeroR y J48</b>	0.9772	0.813
<b>ZeroR y NaiveBayes</b>	0.2113	0.1692
<b>ZeroR y Function SMO</b>	0.03125	0.004545

**Tabla :** Resultados de cada test representado en p-valor

# Dejando un hablante fuera promediando los atributos

Hablantes equilibrados: 8 Buenos Aires, 8 Córdoba

 Hablante para train  Hablante para test









































	<i>Número de hablante</i>										
	1	2	3	4	5	6	7	...	14	15	16
Fold 1								...			
Fold 2								...			
Fold 3								...			
				...							
Fold 16								...			

Tabla : Esquema de cross-validation



# Dejando un hablante fuera promediando los atributos

Juntamos los atributos de cada hablante de la siguiente forma.

Atributos		A1	A2	A3	...	AN
<b>Hablante 1</b>	<b>Audio1</b>	1	?	2		2
	<b>Audio2</b>	?	?	1	...	?
	<b>Audio3</b>	2	?	3		?
<b>Hablante 2</b>	<b>Audio1</b>	1	?	?	...	?
	<b>Audio2</b>	1	2	?		?

Tabla :

esto pasaría a:

Atributos		A1	A2	A3	...	AN
<b>Hablante 1</b>	<b>Audio1</b>	<b>1.5</b>	<b>?</b>	<b>1.667</b>	...	<b>2</b>
<b>Hablante 2</b>	<b>Audio1</b>	<b>1</b>	<b>2</b>	<b>?</b>	...	<b>?</b>

Tabla :

# Dejando un hablante fuera promediando los atributos

	<b>ZeroR</b>	<b>JRip</b>	<b>J48</b>	<b>Function SMO</b>	<b>NaiveBayes</b>
<b>Promedio</b>	53.33	60	60	93.33	80

Tabla : Clasificación correcta en porcentaje

	<b>Student Test</b>	<b>Wilcoxon Test</b>
<b>ZeroR y JRip</b>		
<b>ZeroR y J48</b>		
<b>ZeroR y NaiveBayes</b>		
<b>ZeroR y Function SMO</b>		

Tabla : Resultados de cada test representado en p-valor

# Dejando un hablante fuera promediando los atributos desconocidos

Atributos		A1	A2	A3	...	AN
<b>Hablante 1</b>	<b>Audio1</b>	1	?	2		2
	<b>Audio2</b>	?	?	1	...	?
	<b>Audio3</b>	2	?	3		?
<b>Hablante 2</b>	<b>Audio1</b>	1	?	?	...	?
	<b>Audio2</b>	1	2	?		?

Tabla :

Se cambia a ...

Atributos		A1	A2	A3	...	AN
<b>Hablante 1</b>	<b>Audio1</b>	1	?	2		2
	<b>Audio2</b>	<b>1.5</b>	?	1	...	<b>2</b>
	<b>Audio3</b>	2	?	3		<b>2</b>
<b>Hablante 2</b>	<b>Audio1</b>	1	<b>2</b>	?	...	?
	<b>Audio2</b>	1	2	?		?

Tabla :

# Dejando un hablante fuera promediando los atributos desconocidos

	ZeroR	JRip	J48	Function SMO	NaiveBayes
Fold 1					
Promedio					

Tabla : Clasificación correcta en porcentaje