# Métricas de mimetización acústico-prosódica en hablantes y su relación con rasgos sociales de diálogos

Juan Manuel Pérez

23 de marzo de 2016

# Sistemas de diálogo Humano-Computadora

Ejemplos de "falta de naturalidad"

Relativamente bien en lo lingüístico, pero mal en las emociones, actitudes, intenciones, etc.

### Ejemplos:

- Sistemas telefónicos
- Apple Siri
- Google Now
- 4 Amazon Echo
- Microsoft Cortana
- HAL 9000



### Mimetización

- Mimetización: Fenómeno insconsciente que se manifiesta a través de la adaptación de los hablantes. Fuertemente emparentada con el sentimiento de empatía.
- Objetivo del trabajo: Explorar, refinar y validar una métrica de la mimetización acústico-prosódica.

### **Temario**

- Introducción
  - Prosodia
  - Mimetización
- 2 Trabajo Previo
  - Otras Métricas
  - Método TAMA
- Materiales y Métodos
  - Corpus
  - Modificaciones a TAMA
  - Análisis de la relación con variables sociales
- 4 Resultados
- Conclusiones

# Prosodia ¿Qué es?

- El "cómo" decimos las cosas, a diferencia del "qué"
- Parte fundamental de la comunicación oral.
- Características que la definen: acentuación, velocidad, tono, ritmo, volumen.
- Una de las principales fallas en los sistemas humano-computadoras hoy día
- Transmite emociones, actitudes, y también desambigua.
- Ej. de desambiguación: "Hola, qué pasa?"

### Mimetización

¿Qué es?



- También conocida como entrainment, convergencia, efecto camaleón, etc.
- Adaptación que ocurre entre hablantes a varios niveles: sintáctico, prosódico, en las posturas, etc.
- Fenómeno ubicuo e inconsciente en la comunicación humana
- Para el presente trabajo, sólo nos interesa medir la mimetización acústico-prosódica sobre variables como el tono o pitch, volumen, calidad de habla, y otras.

### Mimetización

#### ¿Y cómo la medimos?

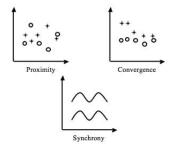
- La definición de mimetización hasta acá vista es muy subjetiva ¿Cómo definimos una medida para esto?
- Vamos a explorar una métrica definida en trabajos anteriores, pulirla un poco, y verificar que efectivamente capture ciertas características del mimetización.
- ¿Cómo? Aplicándola a un corpus con anotaciones sociales, y verificando la relación entre las percepciones sociales y la métrica del mimetización

### **Temario**

- Introducción
  - Prosodia
  - Mimetización
- 2 Trabajo Previo
  - Otras Métricas
  - Método TAMA
- Materiales y Métodos
  - Corpus
  - Modificaciones a TAMA
  - Análisis de la relación con variables sociales
- 4 Resultados
- Conclusiones

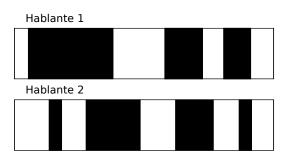
### Otras métricas

#### de mimetización acústico-prosódica



- Algunas métricas captan sólo la convergencia de los factores
- Otras capturan la sincronía: la variación turno a turno.
- Algunas métricas requieren de anotaciones manuales sobre las conversaciones, por ejemplo patrones de entonación.

## Problema del alineamiento de tiempo



Uno de los problemas que tenemos a la hora de construir métricas de mimetización

- ¿Cómo comparamos los diferentes turnos de una conversación?
- Comparar uno a uno es un enfoque simplista y no representativo de la realidad

### Método TAMA

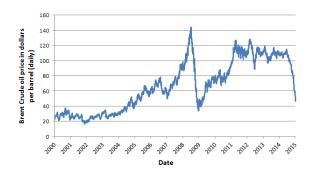
#### Time Aligned Moving Average

- Introducido en Kousidis et al (2008) <sup>1</sup>
- Construímos en primer lugar series de tiempo para cada uno de los hablantes, dada una variable acústico/prosódica.
- Mimetización se define como una función de estas dos series.

<sup>&</sup>lt;sup>1</sup>Spyros Kousidis, David Dorran, Ciaran McDonnell, and Eugene Coyle. Times series analysis of acoustic feature convergence in human dialogues. In Proceedings of Interspeech. 2008.

### Series de Tiempo

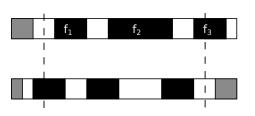
¿Qué es una serie de tiempo?



- En términos coloquiales, una serie de tiempo es una colección de datos temporales.
- El tipo de análisis que vamos a efectuar es propio de Economía y Ciencias de la Atmósfera.
- ¡Mucho más manejables que una sucesión de turnos!

### Método TAMA

#### Cómo construyo la serie de tiempo (dada una variable a-p)



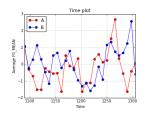
- Segmento de habla considerado Segmento de habla no utilizado
- Silencio Límite de Ventana

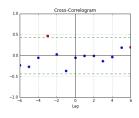
- Partimos la conversación en ventanas solapadas.
- Para cada ventana: calculamos un promedio ponderado del valor de la variable acústico-prosódica en cada segmento de habla

$$\mu = \sum_{i=1}^{N} f_i \frac{d_i}{D} \text{ con } D = \sum_{i=1}^{N} d_i$$

### Método TAMA

#### ¿Y el mimetización?





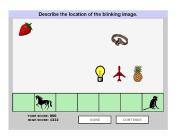
- Ya tenemos la serie de tiempo
- ¿Cómo calculamos la mimetización?
- Función de correlación cruzada: mide la influencia de una serie sobre otra.
- Similar a la correlación, pero aplicando un desplazamiento en alguna de las dos series.
- Los valores significativos de esta son los que se consideran los valores de mimetización (si es que los hay)
- Si el valor es positivo, lo consideramos mimetización.
- Si el valor es negativo, lo consideramos antimimetización.

### **Temario**

- Introducción
  - Prosodia
  - Mimetización
- 2 Trabajo Previo
  - Otras Métricas
  - Método TAMA
- Materiales y Métodos
  - Corpus
  - Modificaciones a TAMA
  - Análisis de la relación con variables sociales
- 4 Resultados
- Conclusiones

# Columbia Games Corpus

#### Descripción



- Desarrollado por Agustín Gravano para su tesis doctoral.
- Corpus de conversaciones diádicas en Inglés Norteamericano
- 12 sesiones con 14 tareas/juegos cada una.
- En cada sesión, se sentó a dos participantes en una cabina profesional de grabación, y una cortina opaca colgando entre ellos para evitar la comunicación visual.
- Los participantes contaron con computadoras a través de las cuales interactuaban mediante juegos.

# Columbia Games Corpus

#### Anotaciones sociales

Cinco anotadores escucharon el audio correspondiente a una tarea del juego y respondieron a varias preguntas sobre los sujetos:

Nombre	Pregunta
contributes-to-completion	¿el sujeto contribuye para el éxito del equipo?
engaged-with-game	¿el sujeto parece comprometido con el juego?
making-self-clear	¿el sujeto se expresa correctamente?
planning-what-to-say	¿el sujeto piensa lo que va a decir?
gives-encouragement	¿el sujeto alienta a su compañero?
difficult-for-partner-to-speak	¿el sujeto le hace difícil hablar a su compañero?
bored-with-game	¿el sujeto está aburrido con el juego?
dislikes-partner	¿al sujeto no le agrada su compañero?

De cada una de estas preguntas obtenemos un puntaje de 0 a 5, para cada hablante de cada tarea.

### Extracción de features acústico-prosódicas

Usando el software Praat  $^2$  se extrajeron las variables acústico-prosódicas para cada segmento de habla

Variable	Descripción					
F0 Mean	Valor medio de la frecuencia fundamenta					
F0 Max	Valor máximo de la frecuencia fundamenta					
Int Mean	Valor medio de la intensidad					
Int Max	Valor máximo de la intensidad					
NHR	Noise-to-harmonics ratio					
Shimmer	Shimmer medido					
Jitter	Jitter medido					
Sílabas/seg	Cantidad de sílabas por segundo					
Fonemas/seg	Cantidad de fonemas por segundo					

<sup>&</sup>lt;sup>2</sup>http://www.fon.hum.uva.nl/praat/

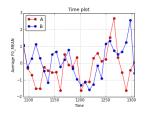
#### **TAMA**

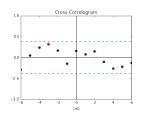
#### Nuestras modificaciones

- Usamos un step de 8s y un tamaño de ventana de 16s, manteniendo el solapamiento del 50 %.
- A diferencia del trabajo original de Kousidis, utilizamos series con datos faltantes.
- Pero sólo nos quedamos con aquellas que tengan 5 o más puntos definidos.

#### **TAMA**

#### Nuestra medida de mimetización





Definimos dos medidas de mimetización

$$\mathcal{E}_{AB}^{(1)} = r_{s}$$
 con s maximizando  $|r_{k}|, k <= 0$   
 $\mathcal{E}_{BA}^{(1)} = r_{s}$  con s maximizando  $|r_{k}|, k >= 0$   
 $\mathcal{E}_{XY}^{(2)} = |\mathcal{E}_{XY}^{(1)}|$ 

Segunda métrica motivada por estudios sobre la antimimetización. Healey et al (2014) sugiere que puede ser una conducta de adaptación cooperativa.

Levitan et al (2015) da más indicios en esa dirección.

## Mimetización y relación con variables sociales

- Llegado a este punto, ya tenemos medidas de mimetización por cada tarea.
- Tambien tenemos las anotaciones sociales
- ¿Cómo analizamos su relación?

### Mimetización y relación con variables sociales

Análisis de Regresión

Para analizar la relación entre las variables sociales (V) y nuestras medidas de *mimetización*  $(\mathcal{E})$ , planteamos un modelo de regresión lineal.

$$V_i \sim \beta_1 + \beta_2 \mathcal{E}_i \tag{1}$$

Nuestra hipótesis es

- **①** Si V es una variable de carácter positivo, entonces  $\beta_2 > 0$
- ② Si V es una variable de carácter negativo, entonces  $\beta_2 < 0$

### **Temario**

- Introducción
  - Prosodia
  - Mimetización
- 2 Trabajo Previo
  - Otras Métricas
  - Método TAMA
- Materiales y Métodos
  - Corpus
  - Modificaciones a TAMA
  - Análisis de la relación con variables sociales
- Resultados
- Conclusiones

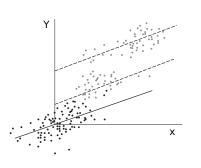
# Resultados de Regresión Agrupada

Regresión Agrupada: Hacemos regresión lineal sobre todos los datos, independientemente de su origen.

- lacktriangledown Para  $\mathcal{E}_{AB}^{(1)}$  no hubo casi resultados significativos
- 2 Para  $\mathcal{E}_{AB}^{(2)}$  hubo resultados significativos, pero pocos

Para analizar mejor los resultados y eliminar las variables no medidas, efectuamos un análisis de Efectos Fijos.

# Regresión Lineal con Efectos Fijos



- Modelo agrupado: niega la posibilidad de heterogeneidad por cada sujeto
- Modelo efectos fijos: heterogeneidad no observada constante en el tiempo para cada sujeto.
- Varias formas equivalentes de calcularlo: Dummy Variable, Within Group.
- En nuestro caso: un sujeto es un hablante dentro de una sesión particular.

# Regresión Lineal con Efectos Fijos

Resultados

	Int Max	Int Mean	F0 Mean	F0 Max	NHR	Fon/seg	Sil/seg	SHIMMER	JITTER
contributes		+	+++	+	++				
clear	+++		+		+++		+		
engaged		+	+++						+
planning									
encourages								+	
difficult		++				_			
bored					+				
dislikes									

Tabla que resume los resultados significativos del análisis con efectos fijos. + indica un valor positivo con significancia  $p<0,\!10,\ ++$  con  $p<0,\!05$  y +++ con  $p<0,\!01.$  Análogamente para los valores negativos.

# Regresión Lineal con Efectos Fijos

#### Resultados

- ① Casi todas las variables acústico-prosódicas poseen al menos un valor significativo de  $\widehat{\beta_2}$ , destacándose *Int Mean*, *NHR* y *F0 Mean*.
- 2 Las variables sociales positivas se relacionan de manera positiva con la métrica de mimetización, en aquellos casos significantes.
- Análogamente ocurre con las variables de connotación negativa, aunque de manera menos clara.
- La métrica se comporta de manera consistente a medidas de mimetización acústico-prosódicas desarrolladas en otros trabajos, como en Gravano et al (2015) utilizando patrones entonacionales. 3

<sup>&</sup>lt;sup>3</sup>Agustin Gravano, Stefan Benus, Rivka Levitan, and Julia Hirschberg. Backward mimicry and forward influence in prosodic contour choice in standard american english 2015

### **Temario**

- Introducción
  - Prosodia
  - Mimetización
- 2 Trabajo Previo
  - Otras Métricas
  - Método TAMA
- Materiales y Métodos
  - Corpus
  - Modificaciones a TAMA
  - Análisis de la relación con variables sociales
- 4 Resultados
- 6 Conclusiones

### Conclusiones

- Desarrollo de métrica de mimetización automática a partir de conversaciones transcritas.
- Indicios de validación de la métrica introducida por Kousidis et al en un corpus orientado a tareas.
- Más indicios sobre la prevalencia y característica positiva del antimimetización

## Trabajo a futuro

- Reproducir experimentos sobre otros conjuntos de datos, por ejemplo Switchboard Corpus. <sup>4</sup>
- Chequear filtro de prewhitening
- Análisis multivariado de las variables acústico-prosódicas para construir nuevas métricas de mimetización

<sup>&</sup>lt;sup>4</sup>https://catalog.ldc.upenn.edu/LDC97S62

# ¿Preguntas?