

# Propuesta de Tesis

Juan Manuel Pérez

4 de mayo de 2015

## Resumen

El *entrainment* (mimetización) es un fenómeno inconsciente que se manifiesta a través de la adaptación de posturas, forma de hablar, gestos faciales y otros comportamientos entre dos o más interactores. A su vez, la ocurrencia de esta mimetización está fuertemente emparentada con el sentimiento de empatía y compenetración entre los participantes.

En esta tesis, nos proponemos explorar una técnica algorítmica para detectar el *entrainment* entre variables prosódicas de dos personas. Esta técnica nos permitirá determinar si existe o no convergencia para ciertos parámetros, y ver como está esto correlacionado con variables sociales tales como la empatía, la compenetración con la tarea, y otras.

## Introducción

Los sistemas de diálogo humano-computadora son cada vez más frecuentes, y sus aplicaciones comprenden una amplia gama de rubros: desde aplicaciones móviles, motores de búsqueda, juegos, o tecnologías de asistencia para ancianos y discapacitados.

Si bien es cierto que estos sistemas logran captar la dimensión lingüística de la comunicación humana, tienen un déficit importante a la hora de procesar y transmitir el aspecto superestructural de la comunicación, que radica en el intercambio de afecto, emociones, actitudes y otras intenciones de los participantes. La habilidad de los participantes de poder expresar, comprender, y reaccionar de acuerdo a estas señales sociales es necesaria para el entendimiento mutuo y una comunicación exitosa.

Un aspecto particular de la comunicación es el fenómeno de *entrainment* (arrastré, mimetización, efecto camaleón), que comprende la adaptación inconsciente de las variables acústicas/prosódicas(a/p) (por ejemplo, el tono de la voz, la velocidad del habla, etc) de manera dinámica en el transcurso de una o varias interacciones. Este fenómeno ha sido introducido por *Brennan et al*[1] en 1996, y se ha observado que la convergencia de los participantes en estas variables ocurre en conjunto con una interacción más fluida y un mayor sentimiento de simpatía por sus interlocutores [2].

Poder medir esta mimetización de los interlocutores no es una tarea fácil, sin embargo. En primer lugar, un diálogo no es una sucesión de turnos, sino que es una serie de tiempo dinámica, llena de interrupciones. Más aún, la mimetización no tiene un carácter instantáneo, sino que se sucede a lo largo de la interacción entre los participantes. Estos factores dificultan ostensiblemente poder modelar este fenómeno.

## Objetivo de la tesis

El objetivo del presente trabajo es verificar un método para medir la mimetización entre dos interlocutores sobre un corpus etiquetado con variables psicosociales, y ver como el *entrainment* está emparejado con la percepción de empatía entre los participantes, compenetración en la charla, y otros.

El método utilizado es *TAMA* (Time-aligned Moving Average), que modela una conversación entre dos interlocutores como series de tiempo. Lo novedoso del trabajo es poder aplicar esta herramienta sobre el corpus “Columbia Games Corpus”. Este corpus de conversaciones posee etiquetas psicosociales, que permitirán corroborar si esta medición de la mimetización de los interlocutores se correlaciona con los rasgos esperados (compenetración en la charla, simpatía con el interlocutor, etc).

## Método

Como ya se mencionó, utilizaremos *TAMA*, introducido en [5]. Consideramos, previo a la utilización de la técnica, a una conversación como una serie de intervalos de elocuciones. Estos intervalos poseen variables acústicas/prosódicas, como pueden ser el tono (frecuencia de la voz), el volumen, y otros.

*TAMA* convierte estas secuencias de elocuciones en dos series de tiempo suavizadas por un promedio móvil, atenuando mayormente los silencios entre turnos. Una vez convertida la conversación en dos charlas, efectuamos análisis estadísticos bi-variados sobre ambas para medir si una de las series “copia” a la otra; es decir, si están mimetizadas.

La mimetización puede ser unidireccional ( $A \rightarrow B$ ) donde una de las conversaciones puede ser predecida por la otra. También puede ocurrir que ambas conversaciones se retroalimenten ( $A \leftrightarrow B$ ).

Finalmente, habiendo determinado si existe *entrainment* en una conversación, veremos si ésto se correlaciona con las etiquetas psico-sociales que tiene el corpus.

## Corpus

El corpus utilizado, *Columbia Games Corpus*[6] consiste de nueve horas de diálogo espontáneo entre pares de sujetos que interactúan en una serie de juegos de computadora. Seis mujeres y siete hombres participaron en la colección del corpus; once de los sujetos volvieron a realizar una conversación con un nuevo compañero.

Cada sesión está compuesta por varias tareas. En cada tarea, los jugadores observan colecciones idénticas de objetos en sus pantallas. Uno de los jugadores (el descriptor) tiene un objeto destacado entre todos los demás, y debe describir la posición de éste a su interlocutor. Los jugadores reciben un puntaje de acuerdo a qué tan bien fueron seguidas las instrucciones dadas por el descriptor. Las parejas completaron 14 de estas tareas, alternando roles. A su vez, los interlocutores fueron separados por cortinas para asegurar que la comunicación fuera exclusivamente oral.

## Referencias

- [1] Brennan: *Lexical entrainment in spontaneous dialog*, 1996
- [2] Chartrand, Bargh: *The chameleon effect: The perception-behavior link and social interaction*, 1999
- [3] De Looze, Scherer, Vaughan, Campbell: *Investigating automatic measurements of prosodic accommodation and its dynamics in social interaction*, 2013
- [4] Kousidis et al: *Towards measuring continuous acoustic feature convergence in unconstrained spoken dialogues*, 2008
- [5] Kousidis, Dorran, McDonnell & Coyle: *Time Series Analysis of Acoustic Feature Convergence in Human Dialogues*, 2009
- [6] Levitan, Gravano, Willson, Benus, Hirschberg & Nenkova, *Acoustic-Prosodic entrainment and social behavior*, 2012
- [7] Chatfield C., *The analysis of time series: an introduction, Third Edition* 1984