

EMPATÍA Y CONVERGENCIA DEL TONO FUNDAMENTAL*

EMPATHY AND PITCH CONVERGENCE

CAMILO QUEZADA

Pontificia Universidad Católica de Chile. Santiago, Chile.
cequezad@uc.cl

JUAN PABLO ROBLEDÓ

Pontificia Universidad Católica de Chile. Santiago, Chile.
jprobled@uc.cl

DOMINGO ROMÁN

Pontificia Universidad Católica de Chile. Santiago, Chile.
dromanm@uc.cl

CARLOS CORNEJO

Pontificia Universidad Católica de Chile. Santiago, Chile.
cca@uc.cl

RESUMEN

En este trabajo se explora la relación entre empatía y convergencia fonética. Se parte de la hipótesis de que la existencia de una actitud empática entre dos personas se traduce en una sincronía observable a nivel de tono fundamental, para lo cual se trabaja con las conversaciones de 27 díadas asignadas a dos condiciones experimentales (definidas como Empática y No-Empática). Se analiza, en algunas preguntas particulares de cada interacción, el despliegue temporal del tono fundamental de ambos participantes para determinar si se produce o no convergencia fonética. Se efectúan análisis estadísticos descriptivos e inferenciales cuyos resultados son congruentes con la manipulación efectuada en cada condición experimental. Los datos obtenidos apoyan la existencia de un correlato entre el F0 y factores disposicionales típicamente relegados en los modelos clásicos de la cognición, pero que están presentes en las interacciones conversacionales cotidianas. Adicionalmente, los datos muestran que el efecto de la empatía es más acentuado en parejas mixtas que en parejas del mismo sexo.

Palabras clave: Convergencia fonética, “entrainment”, empatía, interacción conversacional, sincronía.

* Esta investigación fue financiada por FONDECYT (Chile) a través del proyecto 1100863 asignado al autor para correspondencia Carlos Cornejo. Agradecemos a los miembros del Laboratorio de Investigación y Fenomenología su ayuda en todas las fases de este estudio.

ABSTRACT

This study explored the relation between empathy and phonetic convergence. The working hypothesis was that empathy among two individuals translates into a synchronic process impacting on pitch values. This was studied by recording 27 dyadic interactions assigned to one of two experimental conditions (labeled as Empathic and Non-Empathic). The evolution of pitch in time was observed in some of the questions used to guide the conversations. In each conversation pitch was analyzed for both participants to determine whether phonetic convergence occurred or not. Descriptive and inferential analyses were implemented, the results being consistent with the manipulation devised to distinguish each experimental condition. The final data support the existence of a relation between some phonetic features of speech and dispositional dimensions that although typically overlooked by classical models of cognition are largely present in everyday interactions. Also data support the idea of a stronger empathy effect on mixed-gender dyads than on same-gender dyads.

Keywords: Phonetic convergence, “entrainment”, empathy, conversational interaction, synchrony.

Recibido: 09.04.2012. *Aceptado:* 07.08.2012.

1. INTRODUCCIÓN

La convergencia fonética forma parte de un fenómeno más general que en la literatura ha sido denominado como “entrainment”, “alineación”, “acomodación”, “coordinación” o “convergencia” (Wichmann 2010, Kim y Horton, 2011). Estos términos refieren a un proceso en el cual dos o más hablantes, durante el transcurso de una conversación, se coordinan hasta llegar a “alinearse”, es decir, coincidir en alguna dimensión de interés. Se trata de un fenómeno que a nivel lingüístico se ha estudiado desde varias perspectivas que van desde la duración de la velocidad del habla (Giles y Powesland, 1975) hasta coordinaciones a nivel léxico (Garrod y Anderson, 1987) o sintáctico (Branigan, Pickering, McLean y Cleland, 2007). En términos concretos, se entiende que dos hablantes “convergen” lingüísticamente cuando durante una interacción conversacional comienzan a utilizar con mayor preferencia un repertorio común de lexemas o formas gramaticales, o cuando muestran signos de coincidencia en algún otro rasgo de habla.

La convergencia fonética propiamente tal es una coordinación que se da a nivel fonético-acústico y que se puede observar en rasgos ya sea segmentales, como por ejemplo a nivel de unidades léxicas (Pardo, 2006) o VOT (Fowler, Brown, Saba-

diniy Weihing, 2003) o suprasegmentales (Brazil, 1985). Desde el punto de vista suprasegmental, se la ha estudiado principalmente en relación con los tres parámetros que determinan la prosodia: *pitch*, o tono, intensidad percibida y velocidad de habla. Es precisamente uno de estos parámetros, el *pitch*, el que será objeto de análisis en este trabajo, pues se sabe que influye de manera determinante en las cualidades melódicas de las emisiones, transmite información respecto estados emocionales, género y sobre la personalidad del hablante (Kramer, 1964). Desde el punto de vista de la producción del habla, este parámetro se relaciona directamente con la velocidad de vibración de las cuerdas vocales: a mayor número de vibraciones de las cuerdas vocales por segundo, más alto es el *pitch*. Desde el punto de vista estrictamente acústico, se corresponde con la frecuencia fundamental o f_0 entendida tanto como una relación constante entre los armónicos de la señal periódica propia de los sonidos producidos con vibración de las cuerdas vocales, como con el valor inverso del ciclo de una onda periódica (Johnson, 2003). El aspecto perceptivo de este parámetro es el *pitch*, que permite cualificar como aguda o grave (alta o baja) la altura melódica de un sonido.

Trabajos empíricos han mostrado que si se manipulan los parámetros de la prosodia, se obtienen diferencias en la convergencia percibida (Gregory, Dagan y Webster, 1997). Tomada como variable dependiente en investigaciones experimentales resulta además especialmente dúctil, puesto que permite analizar emisiones diferentes con una alta independencia de la sintaxis y los lexemas empleados en ellas, lo que supone una ventaja importante frente a otros paradigmas (en especial el de repetición léxica o *shadowing*, ampliamente reportado en la literatura, Bulatov, 2009).

Desde el punto de vista teórico, la convergencia fonética ha sido estudiada principalmente dentro de dos modelos. El primero es el de la Teoría de Acomodación Comunicativa de Giles, Nikolas y Coupland (1991), muy afín a propuestas surgidas de la Psicología Social (en particular la Teoría de Identidad Social de Tajfel y Turner, 1979). El segundo es de corte más cognitivo y surge de la línea inaugurada por Pickering y Garrod (2004). Típicamente, los estudios enmarcados dentro del primer modelo enfatizan el rol social del lenguaje como una forma de buscar aceptación dentro de un grupo particular, o como un medio que permite diferenciarse de otro grupo social. Un concepto importante en esta línea es el de “distancia” entre los interlocutores, que se operacionaliza estudiando dimensiones como las variantes dialectales de los hablantes (Kim y Horton, 2011) o la asimetría en el estatus social de los interlocutores (Gregory y Webster, 1996). Los estudios que siguen la línea más cognitiva del segundo modelo, en cambio, conciben el proceso de convergencia como producto del funcionamiento cognitivo y plantean que se trata de un fenómeno *automático* e inconsciente cuyo origen se halla en la manera en que las personas procesan la información lingüística. En esta línea

teórica, la convergencia (en cualquiera de sus planos lingüísticos) es el resultado natural e inevitable de la manera en que la arquitectura cognitiva humana procesa el lenguaje en general.

Ambos modelos presentan limitaciones importantes. Atribuir a la convergencia fonética un carácter automático implica una reducción mecanicista que descuida la dimensión afectiva y disposicional de los seres humanos. En una conversación efectivamente se “procesa” información, pero hay también numerosos factores pragmáticos, emocionales y contextuales capaces de incidir en la materialidad de las emisiones. Ninguno de ellos tiene cabida en el modelo de Pickering y Garrod (2004). En los trabajos que enfatizan aspectos sociolingüísticos, en cambio, se postula que son dimensiones como la distancia social o la identidad de los hablantes las que mayormente explican la posibilidad de observar convergencia entre interlocutores. Si bien es cierto que variables como el género, el rol social o las variantes dialectales efectivamente condicionan la convergencia fonética (Bulatov, 2009), no es menos cierto que una teoría que atribuya el fenómeno exclusivamente a estas dimensiones está muy cerca de un mecanicismo muy similar al anterior: dadas ciertas características sociales, una persona tenderá a converger (o no) con otra independientemente de cuál sea su estado emocional o disposicional, puesto que las dimensiones sociolingüísticas generalmente estudiadas son *inherentes* a las personas.

El trabajo que aquí se presenta gira en torno a dos premisas. La primera es que la convergencia fonética es un proceso primeramente *corporal*, en la medida en que la vocalización coordinada de emisiones lingüísticas tiene un componente fisiológico ineludible. De ahí que se considere adecuado enmarcarla como parte de lo que Nagaoka, Komori y Yoshikawa (2007) denominan *sincronía corporeizada*. En esta perspectiva, cuando dos hablantes se coordinan la sincronía que puedan llegar a establecer entre ellos se puede dar en múltiples niveles corporales, lo que además de rasgos fonéticos incluye sincronía en movimientos gestuales, expresiones faciales, rítmica entonacional, imitación estilística, postura, entre otras dimensiones. Vista así, la convergencia fonética es parte de un proceso más general de sincronía que puede ser estudiado atendiendo a diferentes marcadores fisiológicos. La segunda premisa es que este fenómeno de sincronía corporeizada efectivamente se puede ver favorecido o entorpecido por el grado de distancia generado por rasgos sociolingüísticos inherentes a los hablantes, pero además *covaría* junto a dimensiones psicológicas. Lejos de ser automático y estar determinado exclusivamente por los algoritmos que regularían el procesamiento de diferentes clases de información lingüística (como plantean Pickering y Garrod, 2004), el proceso de sincronía corporeizada es un proceso holístico ligado a constructos psicológicos que influyen en la manera en que las personas interactúan en su vida cotidiana. Uno de estos constructos psicológicos es la empatía.

2. LA NOCIÓN DE EMPATÍA EN PSICOLOGÍA

La empatía ha sido abordada en la psicología recurriendo a diferentes metodologías y desde diversas perspectivas teóricas (Fernández-Pinto, López-Pérez y Márquez, 2008). Desde la psicología social se ha estudiado su vínculo con el comportamiento altruista, es decir, con conductas de ayuda (Batson, Eklund, Chermok, Hoyt y Ortiz, 2007; Stocks, Lishner y Decker, 2009). En la psicología clínica se ha observado su relación con el bienestar psicológico (Angus y Kagan, 2007). Más recientemente, en las ciencias cognitivas se ha explorado el rol de la empatía en la promoción de las interacciones sociales (Decety e Ickes, 2011). Dada esta multiplicidad de enfoques, no resulta sorprendente hallar una considerable variedad de definiciones para el concepto de empatía. Sin embargo, dentro de tal variedad es posible encontrar, al igual que en fonética, distintos niveles de consideración respecto de los aspectos corpóreo e interaccional del fenómeno.

Clásicamente, la empatía ha sido definida como un proceso mental, un estado de conciencia unipersonal. Con frecuencia se estudia su dimensión cognitiva o afectiva, pero también se ha estudiado su relación con la personalidad y con la situación contextual en la que emerge. El enfoque cognitivo se ha servido de herramientas conceptuales como la *toma de perspectiva* (Håkansson, 2006) y la *teoría de la mente* (Gallagher y Frith, 2003). Los enfoques afectivos se caracterizan por poner en segundo plano los estados mentales y destacar los componentes emocionales del fenómeno (Hoffman, 1987). Algunos autores han optado por abordar la empatía como un derivado del ámbito de la personalidad (Davis, Mitchell, Hall y Lothert, 1999) ligado a características más bien estables del individuo. Por otro lado, también se ha abordado el aspecto interaccional y contextual del fenómeno (Batson, Lishner, Cook y Sawyer, 2005; Stocks *et al.*, 2009). Este es el caso de la Hipótesis Empatía-Altruismo, en la que la empatía “refiere a una respuesta emocional orientada hacia un otro, congruente con el bienestar percibido de éste, que resulta de adoptar la perspectiva (por ejemplo, imaginar los sentimientos o pensamientos) de una persona en clara situación de necesidad” (Stocks *et al.*, 2009: 649). Las investigaciones enmarcadas en esta teoría conciben la empatía en términos básicamente representacionales (lo que involucra pensamiento consciente). Así, se utilizan nociones como toma de perspectiva, predicción, atribución, imaginación de los sentimientos del otro, etc. A pesar del intento por asir el carácter contextual e interaccional de la empatía, finalmente el fenómeno es concebido y registrado experimentalmente a través del esfuerzo imaginario de un solo sujeto fuera de contexto, a través de cuestionarios referidos en su mayoría a situaciones hipotéticas (Spreng, McKinnon, Mar y Levine, 2009), algo parecido a lo que ocurre en otras teorías.

Más contemporáneamente, han surgido propuestas que no solo incluyen el rol de procesos psicológicos intangibles, sino también el rol del cuerpo en el estudio

de la empatía. Una aproximación clásica es la basada en el concepto de *proyección sentimental* (Lipps, 1923). Se trata de un proceso mediante el cual un hablante percibe al otro de manera holística, lo que incluye sus *movimientos expresivos* (vocalizaciones, gestos, etc.). Esto genera una vivencia empática más corporalmente amplia que la producida por el plano meramente afectivo. Algunos de estos movimientos son los *sonidos afectivos* (énfasis vocales, tono de voz, ritmo del habla, etc.), que constituirían “la primera materia y ocasión para la reconstrucción de la personalidad ajena” (103). En la misma línea y de manera más reciente, algunos estudiosos del campo de la psicología social han propuesto el concepto de *contagio emocional*, entendido como la tendencia a automáticamente hacer mímica y sincronizar expresiones faciales, vocalizaciones, posturas y movimientos con los de otra persona y, consecuentemente, converger emocionalmente con esta (Hatfield, Rapson y Le, 2009). Una primera hipótesis es que el contagio emocional se produciría a través de la imitación o mímica: la sincronización con los movimientos del otro. Una segunda hipótesis es la retroalimentación: la experiencia emocional es ajustada neuronal (ver también Iacoboni, 2005), motora, o conscientemente en relación al otro. La Hipótesis de Feedback Vocal (Hatfield, Hsee, Costello, Schalekamp y Denney, 1995) está en consonancia con esta última opción. Las emociones estarían conectadas a patrones específicos de entonación vocal que se ajustarían en relación a los emitidos por el interlocutor.

Si se considera que la mayoría de las definiciones clásicas de empatía incluyen la idea de “imaginar” los pensamientos, o incluso “imaginar” los sentimientos del otro, lo que lleva el problema a un nivel representacional, resulta muy relevante el hecho de que tanto la idea de proyección sentimental descrita por Lipps como la de contagio emocional aluden a un enfoque corporeizado y por tanto preconsciente o involuntario. Ambas teorías postulan que la empatía no es solo un fenómeno mental o representacional, sino que comprende un involucramiento físico de todo el organismo durante un encuentro empático. Asimismo, desde la Hipótesis de Feedback Vocal (Hatfield *et al.*, 1995), la interacción de sonidos vocales expresivos entre dos personas en un contexto real parece estar involucrada en vivencias tanto internas como potencialmente externas. En suma, existen razones psicológicas para pensar que la generación de sonidos vocales de dos interactuantes estaría directamente relacionada con lo que sucede entre ellos a nivel psicológico durante la interacción. Dicho de otro modo: Desde un enfoque de la empatía como fenómeno interaccional, corporeizado, resulta factible esperar que en el experimentar el mundo como el otro (Bohart y Greenberg, 1997) de la empatía se diesen fenómenos observables a nivel vocal, es decir fonético. Por lo tanto, sería esperable encontrar un vínculo entre convergencia fonética y empatía. Al ser un fenómeno preconsciente, corporal y espontáneo, la convergencia fonética constituiría una herramienta para acercarse experiencialmente a la vivencia del otro.

De ahí que el propósito principal de esta investigación sea explorar el vínculo

entre empatía y la convergencia fonética, concebida como un indicador particular de sincronía corporeizada entre hablantes en una interacción cara a cara. El diseño experimental presentado a continuación muestra cómo se implementó la hipótesis.

3. METODOLOGÍA

3.1. Participantes

La muestra consistió en 54 estudiantes de la Pontificia Universidad Católica de Chile entre 18 y 28 años ($\bar{x} = 22,2$ años). Los estudiantes participaron voluntariamente y fueron convocados para un estudio interaccional con la idea de que conversaran con algún otro estudiante. Se fijó como indispensable el requisito de que los participantes no se conocieran entre sí, algo que fue observado y cumplido en todas las díadas formadas. La decisión de incluir en la investigación solamente alumnos del mismo plantel universitario permitió controlar la variable diastrática a fin de observar con más claridad el efecto de la manipulación de la variable de interés.

Los sujetos fueron asignados aleatoriamente a dos condiciones experimentales: la Condición Empática (CE) y la Condición No Empática (CNE). Esto permitió generar 27 díadas. La muestra final analizada consistió en 13 grabaciones para la CE y 14 grabaciones para la CNE. En la CE hubo 4 parejas mixtas y 9 del mismo sexo, mientras que en la CNE la composición fue homogénea (7 parejas mixtas y 7 del mismo sexo). Debido a que desde el punto de vista del sexo los dos grupos mostraban una leve asimetría en su composición, los análisis estadísticos incluyeron un desglose de los efectos que considera a variable sexo en cada díada (ver resultados). El promedio de edad en la CE fue de 23,34 años, mientras que en la CNE fue de 21,14. Es importante señalar que se observaron diferencias estadísticamente significativas en las edades de los participantes en ambos grupos, pero que los resultados obtenidos no mostraron cambios sustantivos luego de efectuado un análisis de covarianza en el que se removieron estadísticamente las diferencias de edad, lo que permite asumir que el impacto de la covariable es menor y no afecta la variable dependiente medida.

3.2. Materiales y procedimientos

Los estudiantes fueron convocados al Laboratorio de Interacción y Fenomenología de la Escuela de Psicología de la Universidad Católica, donde se les pidió que leyeran y firmaran un consentimiento informado aprobado por el comité de ética de CONICYT. Además, antes de comenzar la conversación se les solicitó que con-

testaran una modificación local del cuestionario *Toronto Empathy Questionnaire* (TEQ, Spreng *et al.*, 2009). Este cuestionario ha sido validado como una medida psicométricamente confiable del grado de empatía (entendida como un rasgo de personalidad estable) de los encuestados y se utilizó para descartar eventuales diferencias en los niveles basales de empatía de los sujetos que pudieran sesgar los resultados de la manipulación experimental. Una vez concluido el proceso de registro de las interacciones y procesados los datos del TEQ, se comprobó que los niveles de empatía previos a la interacción no diferían estadísticamente en las dos condiciones generadas ($t(59) = 1.629, p = .109$).

Para guiar la conversación se utilizó una adaptación de las preguntas del instrumento Fast Friends (Aron, A., Melinat, E., Aron, E.N., Vaollone, R. y Bator, R. 1997). En ambas condiciones se estableció la tarea de sacar secuencialmente tarjetas numeradas que contenían 10 preguntas. Cada una de las preguntas debía ser contestada por los dos participantes (Anexo 1). La manipulación que generó las dos condiciones experimentales se basó en las instrucciones que recibieron las díadas de cada grupo. En la CE, los estudiantes recibieron la instrucción de responder a las preguntas de la manera más espontánea posible y que escucharan atentamente a sus interlocutores para tratar de conocerlos lo más que pudiesen. En la CNE, en cambio, justo antes de que comenzara la conversación se entregó a los estudiantes una hoja en la que se podía leer lo siguiente:

Tabla I. Instrucciones a los participantes
(adaptado de Aron *et al.*, 1997).

Éste es un estudio sobre conversaciones naturales. Como Uds. saben, en las conversaciones naturales no siempre las personas dicen la verdad, y por esa razón queremos incluir ese hecho al estudio. Aquí hay cuatro sobres, cada uno con una instrucción diferente sobre cómo deben responder las preguntas. Cada uno de Uds. sacará un sobre distinto. Hay cuatro posibles instrucciones:

- Tendrás que decir la VERDAD en TODAS tus respuestas
- Tendrás que decir la VERDAD en todas tus respuestas EXCEPTO en UNA
- Tendrás que decir la VERDAD sólo en la MITAD de tus respuestas
- Tendrás que decir la VERDAD en SOLO UNA de tus respuestas

A cada uno de Uds. le tocará una de estas instrucciones. No existe la posibilidad que ambos tengan la misma instrucción. ¿De acuerdo? ¿Tienen alguna pregunta sobre el procedimiento?

Saquen una tarjeta cada uno, por favor, SIN QUE EL OTRO LA VEA.

Los cuatro sobres entregados a los participantes contenían en realidad la misma instrucción: decir la verdad en *todas* las respuestas. Esto generaba en cada estudiante la idea de que el otro estaba obligado a no decir la verdad en al menos una de las respuestas a 10 de las preguntas sobre las que se basaría la conversación. Se asumió que esta manipulación causaría una distancia entre los interlocutores, distancia que afectaría la posibilidad de establecer vínculos empáticos y que podría observarse de manera más plena en la pregunta 9, ubicada hacia el final de la conversación, en la que se pedía a los participantes que relataran las circunstancias en las que habían vivido el terremoto que afectó a Chile en febrero de 2010.

En un cuestionario aplicado a todos los participantes en forma individual después de la interacción se incluyó la siguiente pregunta: “¿Qué tan verídicas te parecieron las historias de tu compañero?”. Cada sujeto evaluó en un puntaje que iba de 1 a 7 la veracidad percibida en los relatos de su interlocutor. Una prueba *t* de muestras independientes efectuada con estos datos entregó los siguientes resultados: $t(52) = 3.596$, $p = .001$. En otras palabras, los participantes de cada condición percibieron niveles significativamente *diferentes* de veracidad en sus interlocutores. Como el promedio más alto de veracidad percibida se observó en la CE, es posible concluir que la manipulación logró efectivamente crear, en los participantes de la CNE, la sensación de que el otro no estaba siendo completamente honesto en sus intervenciones.

3.3. Hipótesis

En congruencia con la alta carga emotiva involucrada en la experiencia del terremoto y con la Hipótesis Empatía-Altruismo (Stocks *et al.*, 2009), se esperaba que la pregunta 9 produjera el mayor grado de empatía y, consecuentemente, de mayor convergencia fonética entre los estudiantes. Debido a las instrucciones entregadas en cada condición, se esperaba también que la convergencia fonética reflejada en la frecuencia fundamental se comportara de manera diferente en los dos grupos conformados, con un mayor grado de convergencia en la CE y uno menor en la CNE. En definitiva, la hipótesis sobre la cual se trabajó era que la posibilidad de ser engañado por el otro generaría una falta de empatía entre los hablantes que se vería reflejada en la presencia o ausencia de convergencia fonética, particularmente en la pregunta 9, cuyo carácter era sumamente emotivo y vivencial.

4. RESULTADOS

Se analizaron las preguntas 1 y 2, ubicadas al comienzo de la interacción, y la pregunta 9 para poder tener dos puntos iniciales de análisis que pudieran servir para ver la evolución en el tiempo de los cambios en el pitch. Como las interaccio-

nes fueron registradas con un solo micrófono, el análisis se efectuó identificando todos los segmentos emitidos por cada hablante en cada pregunta y eliminando los segmentos en los que hubiera solapamiento de voces. Este procedimiento permitió generar archivos que contenían todas las emisiones individuales de cada participante.

Una vez obtenidos los archivos individuales correspondientes a cada pregunta y cada sujeto, se procedió a analizarlos con Praat (Boersma y Weenink, 2011) para obtener los valores de frecuencia fundamental. Se creó el objeto pitch correspondiente usando el método automático de autocorrelación y los valores promedio se obtuvieron a partir de etiquetas dispuestas previamente en los enunciados. La extracción de los datos se hizo mediante un *script* ad-hoc. Es sabido que este tipo de análisis automático introduce esporádicamente errores en el cálculo de la frecuencia fundamental que afectan la precisión del valor final obtenido, pero se asumió que estos errores se comportarían aleatoriamente y su influencia se distribuiría de manera similar en todas las díadas de las dos condiciones experimentales. Una vez obtenidos los valores de pitch de los segmentos por pregunta se ponderaron los valores del pitch observado por la duración del segmento (expresada en segundos). Esto permitió compensar la desigualdad de las extensiones de las intervenciones de cada pregunta, puesto que las preguntas 1 y 2 eran introductorias y en general su extensión era mucho menor a la de la pregunta 9. Finalmente, se obtuvo, para cada participante, un valor de pitch ponderado para cada pregunta y un valor total de pitch ponderado que reunía la totalidad de sus emisiones.

Este total se usó luego para efectuar un análisis basado en la metodología de *Time Series Analysis* (TAMA) de Kousidis, Dorran, McDonnell y Coyle (2009). En concreto, lo que se hizo fue normalizar los valores observados de pitch ponderado de cada pregunta dividiéndolos por el valor de pitch total ponderado. Así se convirtió la frecuencia absoluta ponderada (expresada en Hertz) a un valor positivo en el que el 1 representa la coincidencia exacta con el promedio ponderado. Cualquier incremento o disminución respecto del valor 1 en un hablante se puede interpretar en puntos *porcentuales* de ascenso o descenso respecto de su propio promedio total para la interacción (que en este caso correspondía a la suma de las tres preguntas señaladas anteriormente).

Este método permite observar cómo evoluciona el pitch de cada hablante respecto de su propio promedio y cuenta con la ventaja de trabajar con valores normalizados, lo que elimina la variabilidad frecuencial que se observa siempre al comparar los valores de Hertz correspondientes a personas diferentes, sobre todo cuando no son del mismo sexo. De esta manera, se obtuvo un valor normalizado para las emisiones de cada participante en las preguntas 1, 2 y 9. Estos valores fueron los que se utilizaron finalmente para los dos análisis efectuados.

4.1. Análisis descriptivos

El primer análisis, de carácter descriptivo, muestra la evolución en el tiempo de los participantes en las dos condiciones. Para ello se procedió primero a generar dos grupos al interior de cada condición. En otros estudios de convergencia fonética (Kim y Horton, 2011) se han empleado variables étnicas o diatópicas que permitían identificar claramente los grupos comparados. Sin embargo, como el diseño de la muestra no incluyó ninguna variable que permitiera agrupar a los sujetos en cada díada, y como otras variables relacionadas con el género o la edad no resultaban de interés para esta investigación, se optó por graficar las interacciones en función del valor del pitch ponderado de cada participante. Así, en cada díada se rotuló a los participantes como Bajo o Alto, dependiendo de si el promedio ponderado de su pitch era más alto o más bajo que el de su interlocutor. Luego se obtuvo el promedio de los pitch normalizados por pregunta como se ilustra en los Gráficos 1 y 2.

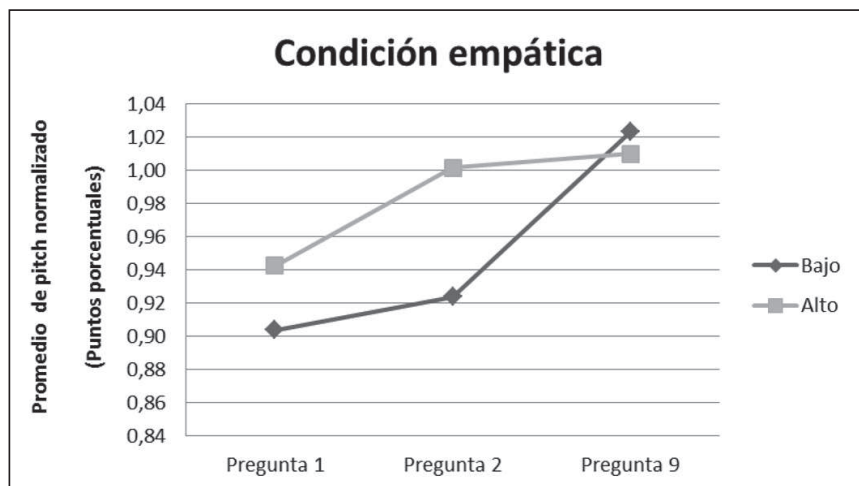


Gráfico 1. Diferencias porcentuales entre preguntas de promedios normalizados de díadas en la Condición Empática.

Promedio de pitch normalizado para las preguntas 1, 2 y 9 para la Condición Empática. Las líneas muestran cómo evolucionan en el tiempo las díadas si en ellas se distingue entre hablantes con un mayor (Alto) o menor (Bajo) promedio total de pitch.

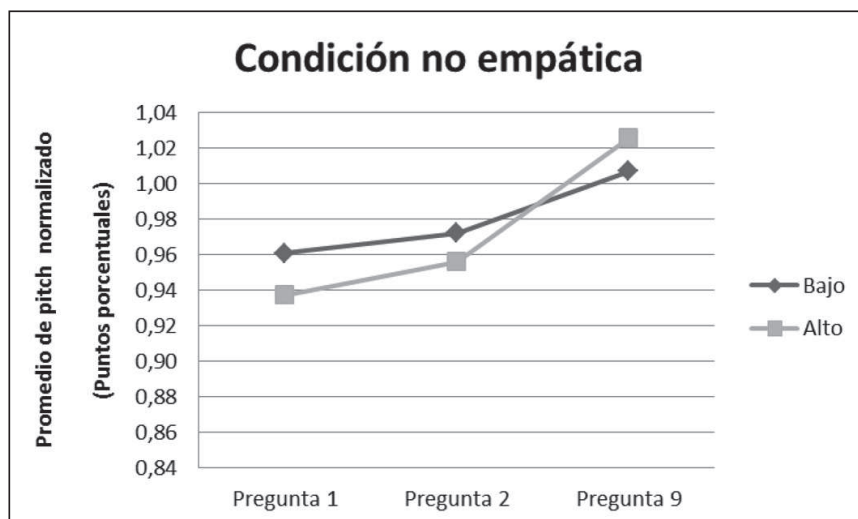


Gráfico 2. Diferencias porcentuales entre preguntas de promedios normalizados de díadas en la Condición No Empática.

Promedio de pitch normalizado para las preguntas 1, 2 y 9 para la Condición No Empática. Las líneas muestran cómo evolucionan en el tiempo las díadas si en ellas se distingue entre hablantes con un mayor (Alto) o menor (Bajo) promedio total de pitch.

En la pregunta de interés (pregunta 9) se puede observar, al comparar los participantes rotulados como Bajo y Alto, una diferencia final de 1,3% en la CE (Gráfico 1) y de 1,8% en la CNE (Gráfico 2), que por supuesto es muy menor. Al revisar la evolución de los promedios en las tres preguntas se observa que en la CE los sujetos de pitch más bajo muestran un incremento muy pronunciado a lo largo de las tres preguntas. En términos porcentuales, su aumento entre la pregunta 1 y la 9 fue de 11,9%, mientras que en la CNE los participantes rotulados como Bajo solo muestran un aumento del 4,6% para el mismo intervalo.

Una de las diferencias más claras entre las dos condiciones está en la manera en que se comportan los sujetos rotulados como Bajo. En la CE, el promedio de los sujetos de pitch más bajo es inferior a los de los de pitch más alto en las preguntas 1 y 2. En la CNE, en cambio, los sujetos de pitch más bajo muestran un promedio superior a los participantes de pitch más alto para las mismas preguntas. Esta diferencia se puede apreciar de manera más clara comparando el comportamiento de los participantes rotulados como Alto en las dos condiciones. En la CE, la diferencia entre la pregunta 1 y la 9 es de 6,7%, mientras que en la CNE es de 8,8%. Resulta llamativo también el hecho de que en la CNE el aumento entre la pregunta 1 y 2 sea de solo 1,8% mientras que en la CE sea de 5,8%.

Analizados de manera conjunta, estos datos sugieren que los participantes de la CE tardaron muy poco en llegar a un nivel de voz natural, algo que demoró mucho más en la CNE. Es importante destacar que gracias a la normalización efectuada el valor 1.0 del eje *y* representa la coincidencia perfecta del pitch ponderado de esa pregunta con el promedio de pitch ponderado total. En otras palabras,

la coincidencia perfecta de cada sujeto respecto de su propio promedio ponderado en todos los segmentos considerados para el análisis. Con esto en mente, el hecho de que en la CE los participantes rotulados como Alto lleguen casi de inmediato al valor 1.0 (su propio promedio en la interacción) puede ser interpretado como una muestra de espontaneidad desde el comienzo mismo de la interacción, algo que no sucede en la CNE. En líneas generales, este primer análisis descriptivo permitió pensar que la manipulación logró generar patrones diferentes en cada condición.

4.2. Análisis inferenciales

Para complementar los resultados anteriores se efectuó primeramente una prueba de ANOVA mixto de 2x3, con un factor intersujetos de dos niveles correspondiente a las condiciones experimentales y un factor intrasujetos de tres niveles correspondiente a las preguntas 1, 2 y 9. Como variable dependiente se tomó la diferencia de pitch normalizados para cada pregunta. Más precisamente, se observaron, en cada día, los valores normalizados para las preguntas 1, 2 y 9 y se efectuó una resta. Esto permitió obtener, para cada pareja de interlocutores, tres valores (correspondientes a las preguntas 1, 2 y 9) que reflejaban una diferencia expresada en términos porcentuales respecto del promedio total ponderado empleado para la normalización. Se trabajó solo con valores positivos (cada diferencia se elevó al cuadrado y luego se obtuvo la raíz) para poder promediar adecuadamente y porque el signo negativo no reflejaba más que la posición de cada hablante en la sustracción, algo que para los efectos de la prueba no resultaba de interés. Así, para la pregunta 1 se obtuvieron 13 valores de diferencia para la CE y 14 para la CNE. Lo mismo con las preguntas 2 y 9. Finalmente, se efectuó la prueba de ANOVA cuyos resultados se reportan a continuación (Gráfico 3).

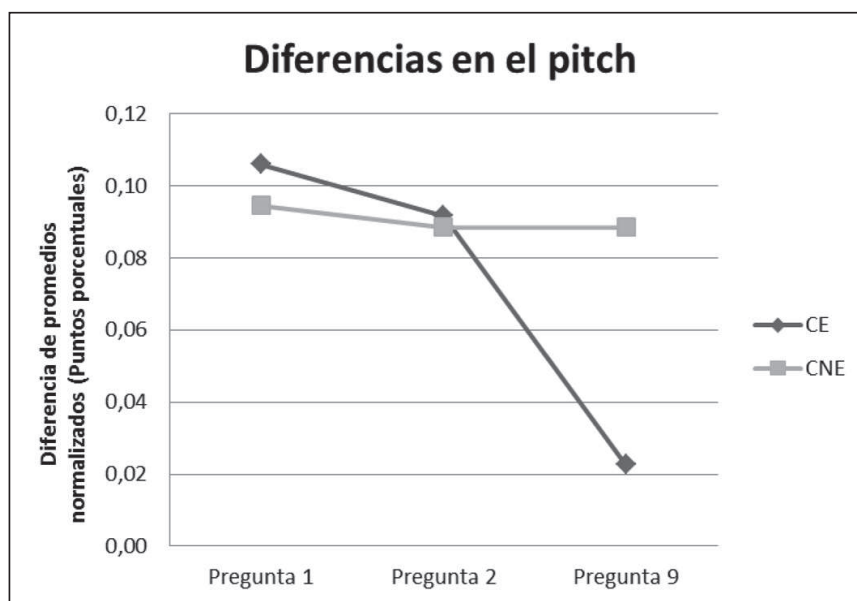


Gráfico 3. Diferencia promedio en el pitch entre ambos interactuantes de cada día en tres preguntas diferentes, en ambas condiciones.

Diferencia promedio en el pitch entre los dos interactuantes de cada díada conversacional en tres momentos subsecuentes de la conversación (en ambas condiciones). Corresponde a un ANOVA mixto con un factor intersujeto de dos niveles (condición empática y condición no-empática) y un factor intrasujeto de tres niveles (preguntas 1, 2 y 9).

La prueba general de ANOVA arrojó, para la variable intrasujetos (las preguntas) el siguiente resultado: $F(1,4) = 2.787$, $p = .092$ (con la corrección de Greenhouse-Geisser). El tamaño de efecto obtenido es de .33 (eta parcial de .01), con una potencia observada de .78. La interacción entre las condiciones experimentales y las preguntas generó los siguientes resultados: $F(1,4) = 2.298$, $p = .13$ (con la corrección de Greenhouse-Geisser), con un tamaño de efecto de 0.3 (eta parcial de .084). A pesar de no ser estadísticamente significativos, estos resultados son bastante buenos desde el punto de vista del tamaño de efecto, lo que permite pensar en que hay ciertamente diferencias entre las condiciones. Es muy probable que una muestra mayor habría permitido llegar a la significación estadística.

Como se puede ver en el gráfico anterior, que muestra los promedios de las diferencias para cada pregunta, tanto la diferencia entre las preguntas como la interacción se basan casi exclusivamente por el descenso que se observa en la CE. Una vez aplicada una prueba de Bonferroni de comparaciones múltiples por pares para observar con mayor detalle las diferencias entre preguntas y grupos, resultó muy evidente que las tendencias son distintas en ambas condiciones. Los valores brutos de los promedios de la CNE prácticamente no varían entre una pregunta y otra y las diferencias entre cada pregunta no son estadísticamente significativas. En la CE, en cambio, se puede ver que el promedio final de las diferencias para la pregunta 9 es de apenas 0,022, mientras que para la CNE este promedio es de 0,088. La diferencia entre las dos condiciones experimentales en la pregunta 9 no alcanza a ser estadísticamente significativa $t(24) = 1.72$, $p = .097$, pero puede ser considerada como una tendencia que muy posiblemente permitiría alcanzar la significación en caso de contar con una muestra más grande. Por otro lado, en la CE la diferencia entre la pregunta 1 y 9 es estadísticamente significativa $t(24) = 2.86$, $p = .024$, tal como la diferencia entre la pregunta 2 y la 9 $t(24) = 5.3$, $p = .003$. Por último, cabe destacar que hay un efecto significativo para la tendencia lineal de los promedios de la variable Pregunta, $F(1,25) = 4.925$, $p = .036$, lo que, dado el comportamiento de los promedios graficados, se debe sin duda a la disminución progresiva y lineal del promedio en la CE.

Para estudiar la influencia de la variable Sexo se distinguió entre díadas del mismo sexo y díadas mixtas en cada condición experimental. Luego, se generó con los mismos datos y la misma variable dependiente una prueba de ANOVA mixto de $2 \times 2 \times 3$, con Condición y Sexo como factores intersujetos y las preguntas 1, 2 y 9 como factor intrasujetos.

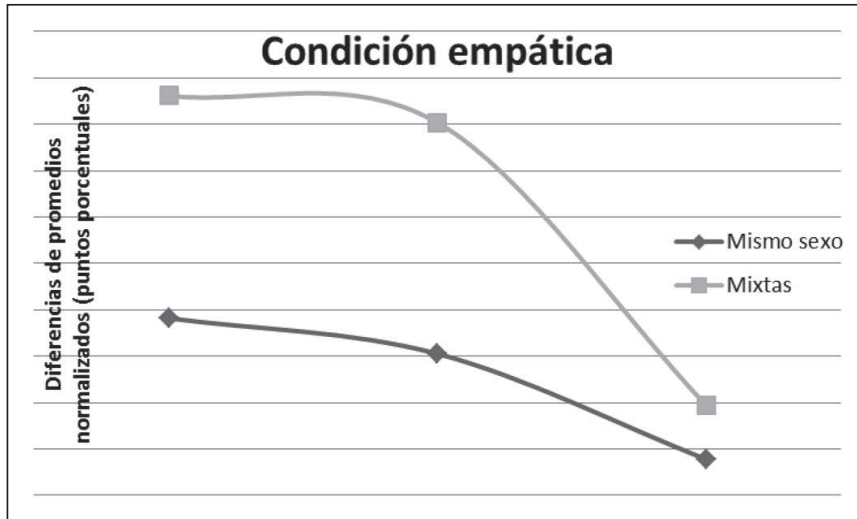


Gráfico 4. Diferencia promedio en el pitch entre los interactuantes de los dos tipos de díada en la condición empática.

La prueba de comparación por pares con ajuste de Bonferroni sólo muestra significación estadística para el contraste de las preguntas 1 y 9 en las díadas mixtas de la CE, $t(23) = 3,26$, $p = .001$.

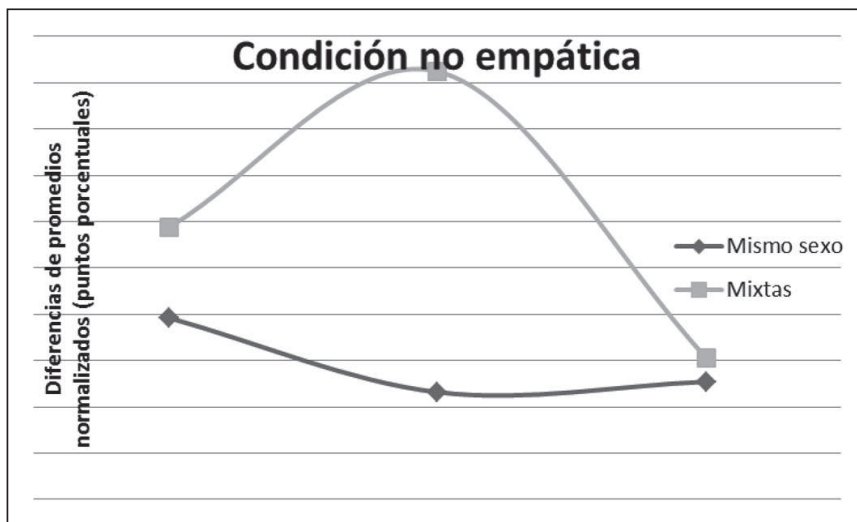


Gráfico 5. Diferencia promedio en el pitch entre los interactuantes de los dos tipos de díada en la condición no empática.

La única diferencia estadísticamente significativa (ajuste de Bonferroni) se observa entre las díadas del mismo sexo y las díadas mixtas en la pregunta 2 de la CNE: $t(23) = 2,1$, $p = .048$.

La prueba general de ANOVA arrojó, para la variable intrasujetos (esto es,

las diferencias normalizadas de pitch de cada díada en las preguntas 1, 2 y 9) el siguiente resultado: $F(1,51) = 5,436$, $p = .015$ (con la corrección de Greenhouse-Geisser). El tamaño de efecto obtenido es de .48 (eta parcial de .191), con una potencia observada de .90. En otras palabras, al incluir la variable Sexo en el modelo se obtiene una alta potencia y un efecto significativo para la variable intrasujetos. Ninguna de las otras interacciones del modelo general resultó significativa estadísticamente.

Al repetir el análisis efectuado en la prueba anterior, es decir, al efectuar una comparación por pares (con ajuste de Bonferroni) de todos los tipos de díadas en las medidas repetidas, se obtiene una única comparación significativa: la que contrasta los promedios de las preguntas 1 y 9 en las díadas mixtas de la CE, $t(23) = 3,26$, $p = .001$. Del resto de las comparaciones, la única que está por debajo de una significación de .01 (aunque sin ser significativa) es la correspondiente a la de las preguntas 2 y 3 para las díadas mixtas de la CNE. Al cambiar la perspectiva y ver las diferencias entre tipos de díada en cada pregunta se obtiene una sola comparación significativa, la correspondiente a la diferencia de promedios entre las díadas del mismo sexo y mixtas en la pregunta 2 de la CNE: $t(23) = 2,1$, $p = .048$. En resumen, como se puede apreciar en los gráficos, el tipo de díada ejerce una influencia sobre la convergencia de los pitch normalizados, pero esta influencia muestra un comportamiento diferente en cada condición. Específicamente, el efecto de la variable empatía es más visible en díadas de distinto sexo que en díadas del mismo sexo.

5. CONCLUSIONES

En primer lugar, los resultados muestran que la manipulación experimental efectuada generó diferencias entre los dos grupos de participantes. Estas diferencias se pueden observar en los distintos patrones observados tanto en el análisis descriptivo como en el correspondiente a la prueba de ANOVA mixto. En las pruebas de ANOVA, como se vio, algunas de las diferencias son estadísticamente significativas y algunas otras no llegan a serlo, pero están muy cerca de los valores críticos típicamente requeridos para la significación estadística. Lo más llamativo de los resultados de la prueba general de ANOVA de 2×3 es la nula evolución de las diferencias entre los hablantes entre cada pregunta analizada. Mientras los sujetos en la CE mostraron una tendencia clara a disminuir las diferencias normalizadas de sus pitch en cada pregunta, los sujetos de la CNE no mostraron ninguna progresión entre cada pregunta. En otras palabras, mientras los sujetos de la CE tendieron a converger en el tiempo, los sujetos en la CNE mantuvieron la misma "distancia" durante los segmentos de la interacción analizados. Esto resulta especialmente llamativo dada la naturaleza altamente emocional de la pregunta 9 y queda además

reforzado por los resultados del ANOVA que incluye la variable Sexo.

Todo esto apoya la hipótesis principal del estudio y permite establecer varias conclusiones. En primer lugar, los resultados muestran la existencia de un vínculo entre un fenómeno psicológico, la empatía, y un fenómeno fonético-acústico. Tradicionalmente, se ha asumido que la empatía es principalmente un constructo psíquico, lo que ha dejado en un segundo plano su aspecto físico o material. Sin embargo, los datos de este estudio muestran cómo la empatía tiene una incidencia directa sobre un fenómeno fonético-acústico (y por tanto corporal) como es la convergencia fonética.

En segundo lugar, los resultados permiten afirmar que la empatía no es exclusivamente un fenómeno unipersonal sino que puede emerger –o no– en las relaciones interpersonales. De ahí que el término “sincronía corporeizada” resulte particularmente preciso para capturar tanto el carácter corporal de la empatía como su aspecto esencialmente interaccional: en las conversaciones cara a cara, la empatía está orientada hacia un otro que está física y sensorialmente presente. Cuando hay sincronía entre hablantes es porque los interlocutores se retroalimentan mutuamente en múltiples niveles corporales. La convergencia fonética es tan solo uno de varios marcadores fisiológicos que pueden dar cuenta de sincronía.

La evidencia del presente estudio apoya la idea de que experimentar sincronías y, consecuentemente, converger emocionalmente con otro está mediado por factores contextuales y psicológicos. Si bien en apariencia los resultados son compatibles con la propuesta del *contagio emocional*, existe una diferencia importante: la convergencia fonética no se produce automáticamente. Situando la convergencia fonética entre los mecanismos hipotetizados en la propuesta del contagio emocional, habría que ubicarla conceptualmente como una alternativa intermedia entre el mero *matching* neuronal sin actividad motriz y la imitación o mímica automática. Constituiría una de las herramientas para acercarse experiencialmente a la vivencia del otro. Asimismo, no es sinónimo de imitación, pues la imitación exacta del pitch de un interlocutor es no solo difícil de lograr sino que muchas veces susceptible de ser considerada como algo ofensivo. De hecho, no se dio en ninguna de las interacciones. La convergencia fonética está más cercana a la idea de retroalimentación; los hablantes muestran cambios en el promedio de su pitch (en Hertz) en la medida en que se ajustan a sus interlocutores.

Por otro lado, estos resultados permiten efectuar una reflexión metodológica que puede contribuir a estudios futuros. Como ya se mencionó, uno de los paradigmas más empleados en la literatura para estudiar la convergencia fonética es el *shadowing*, muchas veces complementando acuerdos interjueces o valoraciones de terceros que evalúan el grado de similitud de dos emisiones. Esta forma de estudiar el fenómeno lo priva de dos componentes esenciales para el despliegue de la empatía: la interacción cara a cara con un interlocutor físicamente presente y la evolución en el tiempo. Los resultados obtenidos permiten afirmar que ambas

dimensiones son centrales para que dos personas puedan mostrar rasgos corporeizados de sincronía.

Por último, los datos apoyan también la idea de que los procesos de convergencia o sincronía están lejos de ser automáticos y que en ellos intervienen dimensiones afectivas, contextuales y psicológicas tradicionalmente relegadas en muchas de las teorizaciones sobre lenguaje e interacción. No hay duda de que las variables típicamente investigadas en los estudios que observan la convergencia fonética (variables como género, edad, estrato social y otras) forman parte de las dimensiones que intervienen en los procesos de interacción humana. La investigación presentada complementa dichos estudios mostrando la influencia de variables contextuales y disposicionales en un tipo de convergencia fonética. En particular es importante destacar que la variable empatía muestra su mayor efecto en díadas mixtas. Ello sugiere que variables disposicionales cooperan con rasgos personales y sociodemográficos.

REFERENCIAS

- Angus, L. y Kagan, F. (2007). Empathic relational bonds and personal agency in psychotherapy: Implications for psychotherapy supervision, practice, and research. *Psychotherapy: Theory, Research, Practice, Training*, 44(4), 371-377.
- Aron, A., Melinat, E., Aron, E.N., Vaollone, R. y Bator, R. (1997). The experimental generation of interpersonal closeness: A procedure and some preliminary findings. *Personality and Social Psychology Bulletin*, 23, 363-377.
- Batson, C., Eklund, J., Chermok, V.L., Hoyt, J.L. y Ortiz, B.G. (2007). An additional antecedent of empathic concern: Valuing the welfare of the person in need. *Journal of Personality and Social Psychology*, 93(1), 65-74.
- Batson, C., Lishner, D.A., Cook, J. y Sawyer, S. (2005). Similarity and Nurturance: Two possible sources of empathy for strangers. *Basic & Applied Social Psychology*, 27(1), 15-25.
- Bohart, A.C. y Greenberg, L.S. (1997). *Empathy reconsidered: New directions in psychotherapy*. Baltimore: United Book Press.
- Boersma, P. y Weenink, D. (2011). Praat: doing phonetics by computer [Programa computacional]. Version 5.3, obtenido en noviembre del 2011 desde <http://www.praat.org/>
- Branigan, H.P., Pickering, M.J., McLean, J.F. y Cleland, A.A. (2007). Participant role and syntactic alignment in dialogue. *Cognition*, 104, 163-197.
- Brazil, D. (1985). Phonology: Intonation in discourse. En T.A. Van Dijk (ed.), *Handbook of discourse analysis II: dimensions of discourse* (pp. 55-75). London: Academic Press.
- Bulatov, D. (2009). The effect of fundamental frequency on phonetic convergen-

- ce. *UC Berkeley Phonology Lab Annual Report*, 404-434.
- Davis, M., Mitchell, K., Hall, J. y Lothert, J. (1999). Empathy, expectations, and situational preferences: personality influences on the decision in volunteer helping behaviors. *Journal of Personality*, 67(3), 469-503.
- Decety, J. e Ickes, W. (2011). *The social neuroscience of empathy*. Cambridge, MA: The MIT Press.
- Fernández-Pinto, I., López-Pérez, B. y Márquez, M. (2008). Empatía: medidas, teorías y aplicaciones en revisión. *Anales de Psicología*, 24(2), 284-298.
- Fowler, C.A., Brown, J., Sabadini, L. y Weihing, J. (2003). Rapid access to speech gestures in perception: Evidence from choice and simple response time tasks. *Journal of Memory and Language*, 49, 396-413.
- Gallagher, H.L y Frith, C.D. (2003). Functional imaging of 'theory of mind'. *Trends Cogn Sci*, 7(2), 77-83.
- Garrod, S. y Anderson, A. (1987). Saying what you mean in dialogue: A study inconceptual and semantic co-ordination. *Cognition*, 27, 181-18.
- Giles, H. y Powesland, P.F. (1975). *Speech styles and social evaluation*. New York: Academic Press.
- Giles, H., Nikolas C. y Coupland, J. (1991). Accommodation theory: communication, context, and consequence. En H. Giles, J. Coupland y N. Coupland (eds.), *Contexts of accommodation* (pp. 1-68). New York: Cambridge University Press.
- Gregory, S.W. y Webster, S. (1996). A nonverbal signal in voices of interview partners effectively predicts communication accommodation and social status perceptions. *Journal of Personality and Social Psychology*, 70, 1231-1240.
- Gregory, S.W., Dagan, K. y Webster, S. (1997). Evaluating the relations between vocal accommodation in conversational partners' fundamental frequencies to perceptions of communication quality. *Journal of Nonverbal Behavior*, 21, 23-43.
- Håkansson, J. (2006). Empathy and viewing the other as a subject. *Scandinavian Journal of Psychology*, 47, 399-409.
- Hatfield, E., Hsee, C., Costello, J., Schalekamp, M. y Denney, C. (1995). The impact of vocal feedback on emotional experience and expression. *Journal of Social Behavior and Personality*, 1495, 10(2), 293-312.
- Hatfield, E., Rapson, R.L. y Le, Y.L. (2009). Emotional contagion and empathy. En J. Decety y W. Ickes (eds.), *The social neuroscience of empathy* (pp. 19-30). Cambridge, MA: The MIT Press.
- Hoffman, M. L. (1987). The contribution of empathy to justice and moral judgment. In N. Eisenberg y J. Strayer (eds.), *Empathy and its development* (pp. 47-80). Cambridge: Cambridge University Press.
- Iacoboni, M. (2005). Understanding others: Imitation, language, and empathy. En S. Hurley y N. Chater, *Perspectives on imitation: From neuroscience to social*

- science, Vol. 1, Mechanisms of imitation and imitation in animals* (pp. 77-101). Cambridge, MA: The MIT Press.
- Johnson, K. (2003). *Acoustic and Auditory Phonetics*, 2ª ed. Oxford: Blackwell.
- Kim, M. y Horton, W.S. (2011). Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology*, 2, 125-156.
- Kousidis, S., Dorran, D., McDonnell, C. y Coyle, E. (2009). *Time series analysis of acoustic feature convergence in human dialogues*. Dublin Institute of Technology, Digital Media Centre Conference papers.
- Kramer, E. (1964). Elimination of verbal cues in judgments of emotion from voice. *Journal of Abnormal and Social Psychology*, 68, 390-396.
- Lipps, T. (1923). *Los fundamentos de la estética*. Madrid: D. Jorro.
- Nagaoka, C., Komori, M. y Yoshikawa, S. (2007). Embodied synchrony in conversation. En T. Nishida (ed.), *Conversational informatics* (pp. 331-351). Chichester, UK: Wiley.
- Pardo, J.S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America*, 119, 2382-2393.
- Pickering, M.J. y Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(2), 169-225.
- Rizzolatti, G. (2005). The mirror neuron system and imitation. En S. Hurley y N. Chater, *Perspectives on imitation: From neuroscience to social science: Vol. 1 Mechanisms of imitation and imitation in animals* (pp. 55-76). Cambridge, MA: The MIT Press.
- Spreng, R.N., McKinnon, M., Mar, R. y Levine, B. (2009) The Toronto Empathy Questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of Personal Assessment*, 91(1), 62-71.
- Stocks, E., Lishner, D. y Decker, S. (2009). Altruism or psychological escape: Why does empathy promote prosocial behavior? *European Journal of Social Psychology*, 39(5), 649-665.
- Tajfel, H. y Turner, J.C. (1979). An integrative theory of inter-group conflict. En S. Worchel y W.G. Austin (eds.), *The social psychology of intergroup relations* (pp. 33-47). Pacific Grove, CA: Brooks & Cole.
- Wichmann, A. (2010). Intonational meaning in institutional settings: The role of syntagmatic relations. *Media*, 5(4), 849-857.

ANEXO 1

Protocolo Fast Friends (modificado)

1. Dime cómo te llamas y cuéntame qué carrera estás estudiando
2. ¿Por qué entraste a estudiar esa carrera?
3. ¿Puedes contarme cómo fue tu primer año en la universidad?
4. ¿Qué te gustaría hacer profesionalmente después de terminar la carrera?
5. ¿Con quiénes vives?, ¿tienes hermanos?, ¿cuántos?
6. ¿Dónde vas a ir de vacaciones este verano?
7. ¿Recuerdas cómo eran tus vacaciones cuando eras niño?
8. ¿Realizas alguna actividad aparte de la universidad? (como por ejemplo algún hobby), ¿qué es lo que te gusta de ella?
9. Cuéntame cómo viviste el terremoto en febrero ¿Dónde estabas?, ¿te asustaste?
10. Si pudieras elegir con libertad, ¿en qué lugar estarías ahora?, ¿haciendo qué?