

LOS SIGNOS DEL CUERPO

Enfoques multimodales de la mente y el lenguaje

Carolina Scotto
Fernando G. Rodríguez
Irene Audisio
(compiladores)

COLECCIÓN UAI – INVESTIGACIÓN

UAI EDITORIAL

teseo 

LOS SIGNOS DEL CUERPO

Carolina Scotto
Fernando G. Rodríguez
Irene Audisio
(compiladores)

Los signos del cuerpo

Enfoques multimodales de la mente y el lenguaje

Colección UAI – Investigación

UAI EDITORIAL

teseo 

Los signos del cuerpo: enfoques multimodales de la mente y el lenguaje / Carolina Scotto ... [et al.]; compilado por Scotto, Carolina; Fernando Gabriel Rodríguez; Irene Audisio. - 1a ed. - Ciudad Autónoma de Buenos Aires: Teseo; Ciudad Autónoma de Buenos Aires: Universidad Abierta Interamericana, 2019. 412 p.; 20 x 13 cm.

ISBN 978-987-723-206-6

1. Filosofía. 2. Psicología. 3. Filosofía del Lenguaje. I. Scotto, Carolina II. Scotto, Carolina, , comp. III. Rodríguez, Fernando Gabriel , comp. IV. Audisio, Irene, comp.

CDD 419

© UAI, Editorial, 2019

© Editorial Teseo, 2019

Teseo - UAI. Colección UAI - Investigación

Buenos Aires, Argentina

Editorial Teseo

Hecho el depósito que previene la ley 11.723

Para sugerencias o comentarios acerca del contenido de esta obra,
escribanos a: info@editorialteseo.com

www.editorialteseo.com

ISBN: 9789877232066

Las opiniones y los contenidos incluidos en esta publicación son responsabilidad exclusiva del/los autor/es.

Autoridades

Rector Emérito: Dr. Edgardo Néstor De Vincenzi

Rector: Dr. Rodolfo De Vincenzi

Vice-Rectora Académica: Dra. Ariana De Vincenzi

Vice-Rector de Gestión y Evaluación:

Dr. Marcelo De Vincenzi

Vice-Rector de Investigación: Dr. Mario Lattuada

Vice-Rector de Extensión Universitaria: Ing. Luis Franchi

Vice-Rector de Administración: Dr. Alfredo Fernández

Decano Facultad de Psicología y Relaciones Humanas:

Lic. Fernando Adrover

Comité editorial

Lic. Juan Fernando ADROVER

Arq. Carlos BOZZOLI

Mg. Osvaldo BARSKY

Dr. Marcos CÓRDOBA

Mg. Roberto CHERJOVSKY

Dra. Ariana DE VINCENZI

Dr. Roberto FERNÁNDEZ

Dr. Fernando GROSSO

Dr. Mario LATTUADA

Dra. Claudia PONS

Dr. Carlos SPECTOR

Los contenidos de los libros de esta colección cuentan con evaluación académica previa a su publicación.

Presentación

La Universidad Abierta Interamericana ha planteado desde su fundación en el año 1995 una filosofía institucional en la que la enseñanza de nivel superior se encuentra integrada estrechamente con actividades de extensión y compromiso con la comunidad, y con la generación de conocimientos que contribuyan al desarrollo de la sociedad, en un marco de apertura y pluralismo de ideas.

En este escenario, la Universidad ha decidido emprender junto a la editorial Teseo una política de publicación de libros con el fin de promover la difusión de los resultados de investigación de los trabajos realizados por sus docentes e investigadores y, a través de ellos, contribuir al debate académico y al tratamiento de problemas relevantes y actuales.

La colección *investigación TESEO – UAI* abarca las distintas áreas del conocimiento, acorde a la diversidad de carreras de grado y posgrado dictadas por la institución académica en sus diferentes sedes territoriales y a partir de sus líneas estratégicas de investigación, que se extiende desde las ciencias médicas y de la salud, pasando por la tecnología informática, hasta las ciencias sociales y humanidades.

El modelo o formato de publicación y difusión elegido para esta colección merece ser destacado por posibilitar un acceso universal a sus contenidos. Además de la modalidad tradicional impresa comercializada en librerías seleccionadas y por nuevos sistemas globales de impresión y envío pago por demanda en distintos continentes, la UAI adhiere a la red internacional de acceso abierto para el conocimiento científico y a lo dispuesto por la Ley nº: 26.899 sobre *Repositorios digitales*.

institucionales de acceso abierto en ciencia y tecnología, sancionada por el Honorable Congreso de la Nación Argentina el 13 de noviembre de 2013, poniendo a disposición del público en forma libre y gratuita la versión digital de sus producciones en el sitio web de la Universidad.

Con esta iniciativa la Universidad Abierta Interamericana ratifica su compromiso con una educación superior que busca en forma constante mejorar su calidad y contribuir al desarrollo de la comunidad nacional e internacional en la que se encuentra inserta.

Dra. Ariadna Guaglianone
Secretaría de Investigación
Universidad Abierta Interamericana

Índice

Introducción	15
Parte I. Multimodalidad y lenguaje	33
1. El lenguaje verbal también es icónico: correspondencias transmodales y simbolismo sonoro	35
<i>Carolina Scotto (UNC - CONICET)</i>	
2. Señalar el tiempo	79
<i>Irene Audisio (CIFFyH - UNC)</i>	
3. Esquematicidad mimética en las metáforas multimodales....	113
<i>Carolina Mahler (CIFFyH - UNC)</i>	
Parte II. Percepción directa de otras mentes.....	143
4. Comportamientos expresivos naturales y contenidos mentales no proposicionales.....	145
<i>Laura Danón (CIFFyH - UNC)</i>	
5. Algunas reflexiones en torno al conocimiento de las otras mentes	177
<i>Daniel Kalpokas (UNC - CONICET)</i>	
6. Percepción dinámica de expresiones faciales y atribución de emociones.....	199
<i>Zoé Sánchez Barbieri (UNC - CONICET)</i>	
Parte III. Desarrollo ontogenético y multimodalidad.....	229
7. Desarrollo semiótico, multimodalidad y cuerpo	231
<i>Fernando G. Rodríguez (UAI)</i>	
8. Desarrollo intersubjetivo y perceptivo.....	271
<i>Mauricio Martínez (UAI)</i>	

9. Competencia social y reconocimiento de emociones a partir de expresiones faciales y lenguaje corporal en sujetos del espectro autista.....	325
<i>Xilenia María Carreras (SECYT - UNC)</i>	
10. Gestos y palabras	363
<i>Alejandra Menti (UNC - CONICET) y Celia Rosemberg (CONICET - CIIPME)</i>	
Sobre los autores	405

Introducción

En las últimas décadas, la consideración del cuerpo en general ha reingresado a los estudios y debates que tienen lugar tanto en filosofía como en psicología, para los tópicos en los que una y otra disciplina convergen, y en los que, desde sus perspectivas diferentes, deberán llegar, si no quizás a una perfecta inteligencia y a un *desideratum* de concordia que es difícil esperar, al menos a un entendimiento que permita fluir el intercambio de información nueva y el de nuevos ángulos para evaluarla. La readmisión del cuerpo como factor decisivo en los problemas de la percepción y de la cognición en general, de la cognición social y la atribución mental, de la formación de conceptos y el desarrollo ontogenético, y finalmente de las estrategias del sujeto en los procesos de semiosis, incluyendo el aprendizaje de la lengua, se debe menos a una acción reparadora, a un ejercicio de justicia con el que quisieran compensarse años de exilio, que a la propia fuerza con que la corporeidad se hizo presente y desafió al cognitivismo y al funcionalismo como ensayos de interpretación de la naturaleza del conocimiento y de la condición humana en lo que ella posee de más genuino. En efecto, se hizo palmario, en ciertas circunstancias, que el modelo computacional, al suponer que la mente era un puro mecanismo de combinación de símbolos, había proscripto un componente medular de aquella explicación por la que se afanaba.

El cuerpo, no el cerebro (versión transitoria de superación del enfoque cognitivista clásico), se manifiesta mediante una variedad de conductas expresivas: expresiones del rostro, de las manos, actitudes posturales y proxémicas, la prosodia de la enunciación, la kinesis, todo ello

revelándose determinante para la elaboración de la noción de espacio circundante, de intersubjetividad, de pensamiento, del aprendizaje y de nuestros sistemas comunicativos -del más simple al más complejo-, competencias que implican de un modo más visible o menos la marca de origen de ser *con el cuerpo*, o bien de ser, lisa y llanamente, *cuerpo*. Por cierto esta marca de origen, el molde de la corporeidad elemental, subsiste en ocasiones, si se trata de las competencias más sofisticadas (las funciones psicológicas llamadas superiores: razonamiento, lenguaje, etc.), como una muy sutil, casi insensible "marca de agua", pero que ha de encontrarse con saber buscar en el sitio apropiado (un sitio de la experiencia psicofísica, no restringido a la esfera mental).

Uno de los arietes con los que el cuerpo ha ingresado nuevamente a escena es la multimodalidad. Originario de los estudios de percepción, el término se ha propagado hasta afirmarse en otros territorios, de una manera particularmente axial en el de la semiosis (repertorio léxico, actos comunicativos), donde se ha convertido de aferente en eferente, pero asimismo en el del pensamiento conceptual-mimético y en sus expresiones lingüísticas, en el de las manifestaciones de estados emocionales, las interacciones cara a cara y largo etcétera. Se ha colado en problemáticas como la atribución intencional, el significado y la comprensión de los comportamientos expresivos, o en la percepción de estados emocionales e intencionales en general. Se puede afirmar, en relación con los estudios psicológicos sobre primera infancia, que allí ha desembarcado para replantear todas las discusiones, dado que buena parte de la tradición tenía por hecho consumado que la percepción, fuente de enlace entre el sujeto y el fenómeno, era en el nacimiento un caos incoordinado de datos unimodales separados. Por consiguiente, la

multimodalidad, una estrategia evolutiva para la coordinación de *inputs* de diferentes fuentes y sustratos y ubicada en el comienzo de la vida humana, tenía por destino conmover las bases y obligar a una exigente, ineluctable revisión de las viejas premisas. Efecto de ello fue que la corporeidad ocupara una posición central en la investigación de los procesos cognitivos devenidos enactivos (esto es, anclados en el ser físico o biofísico, y ecológicamente coordinados con el medio), que el movimiento corporal se convirtiera de accesorio en principal, que las vocalizaciones y los gestos se volvieran blanco de atención como expresiones no tan solo precursoras, sino también preditoras de las futuras habilidades del niño gramatical, que el tiempo vivencial (no el físico-objetivo) ganara importancia y espesor, que el espacio inmediato y tridimensional fuera reconocido como un emergente de la experiencia del acoplamiento del cuerpo y su entorno, y que -por cierto al fin de esta enumeración, pero en verdad al comienzo de todo- la musicalidad presente en las interacciones adulto-bebé permitiera exhumar matices nunca sospechados en la primera intersubjetividad.

Tales perspectivas corporizadas impactaron fuertemente en los estudios sobre la comunicación humana, y provocaron la gradual revisión de algunos de los rasgos que se consideraban definitorios de la singularidad del lenguaje natural. Por una parte, los estudios sobre gestualidad modificaron drásticamente los enfoques ortodoxos basados en distinguir propiedades nucleares abstractas y amodales y la periferia perceptual-motora en la que ocurren las interacciones comunicativas. Bajo el mismo signo, la lingüística cognitiva, con su propuesta de comprender a la cognición conceptual en tanto estructurada como una metáfora conceptual anclada en la experiencia corporal, fortaleció un enfoque corporizado acerca de la

estructuración de los conceptos. En la filosofía de la mente y en la psicología del desarrollo, la cognición social o atribución psicológica sufrió también un giro en el que las expresiones faciales, los rasgos expresivos del comportamiento y el papel de la percepción desplazaron la omnipresencia de los enfoques intelectualistas y fuertemente epistémicos que habían dominado la escena en aquellas dos áreas. Nombres y fechas para el conjunto de estas transformaciones se hallarán en los capítulos de este libro.

En “Los signos del cuerpo: enfoques multimodales de la mente y el lenguaje”, se reúnen los resultados de las investigaciones presentadas y discutidas en el Primer Encuentro sobre *Gestos y Multimodalidad: Lenguaje, emociones y percepción*, realizado en la Facultad de Filosofía y Humanidades de la Universidad Nacional de Córdoba, los días 2 y 3 de agosto de 2018. En dicho encuentro participaron investigadores que trabajan en distintos ámbitos disciplinares, tales como la filosofía de la mente y del lenguaje, la psicología cognitiva del desarrollo infantil, la lingüística cognitiva y comparada, desde una perspectiva teórica pero también interesados en las implicaciones y aplicaciones de distintos enfoques y resultados empíricos.

En “El lenguaje verbal también es icónico: correspondencias transmodales y simbolismo sonoro”, Carolina Scotto desarrolla un argumento empírico basado en la abundante evidencia en la literatura científica reciente que corrobora la *ubicuidad* y la *importancia relativa* en todos los lenguajes verbales del fenómeno icónico conocido como “*simbolismo sonoro*”. Este concluye que el lenguaje verbal no es solo (mayormente) arbitrario, sino también icónico. Analiza, además, cómo dicho fenómeno es una manifestación en el lenguaje de un fenómeno cog-

nitivo más amplio, el que en la literatura neurocognitiva se conoce como “correspondencias transmodales”. Se trata de aquellas percepciones multisensoriales en las cuales estímulos o atributos de esos estímulos, aparentemente no relacionados entre sí, en diferentes modalidades y dimensiones sensoriales, producen una concordancia, conjunción o efecto de congruencia en el contenido de la experiencia. La vinculación entre el simbolismo sonoro y las correspondencias transmodales apoya la propuesta, defendida por la autora, de elaborar una caracterización radicalmente diferente del lenguaje natural de aquella que han ofrecido los enfoques ortodoxos, tanto en la lingüística como en la filosofía del lenguaje, más afín con los enfoques corporizados que con la perspectiva cognitivista clásica. Esta reformulación se basa en un argumento complementario del anterior: si el simbolismo sonoro, como un subtipo lingüístico de iconicidad, es una expresión del papel de las correspondencias transmodales en el lenguaje, entonces el lenguaje debería ser redescrito como un *fenómeno semiótico heterogéneo y multimodal*. Como consecuencia de ello, y dado que en el conocimiento y en el uso del lenguaje están involucradas dimensiones icónicas y multimodales, Scotto concluye que estos rasgos serán también relevantes para explicar una variedad de otros aspectos del lenguaje, tales como el aprendizaje y el procesamiento lingüístico, los usos comunicativos, los orígenes filogenéticos, el desarrollo ontogenético, el cambio lingüístico, entre otros. La reconfiguración del lenguaje que se propone es ejemplificada en este trabajo mediante el examen de algunas importantes comprobaciones empíricas que sugieren cuál podría ser el papel del simbolismo sonoro en un área específica, como es el aprendizaje temprano del lenguaje.

En “Señalar el tiempo: el gesto deíctico desde un enfoque multimodal del lenguaje”, Irene Audisio se ocupa de evaluar el papel de los gestos en relación con el lenguaje verbal humano. Actualmente se admite que los gestos son una parte integral del lenguaje, o que incluso forman un sistema unificado con el habla. Tanto en el origen del lenguaje como en los procesos comunicativos *online*, los gestos manuales muestran una relación dialéctica con las palabras de la que depende el significado de las expresiones. La autora se ocupa, en particular, del empleo de los gestos naturales deícticos, considerados básicos para la comunicación humana, examinando un uso específico del mismo gesto, aquel ejemplificado por su incorporación en la lengua general de la Amazonia, el nheengatú. En esta lengua, el gesto deíctico natural se une al empleo de palabras como un elemento constitutivo del sistema lingüístico. En efecto, se ha identificado un uso del movimiento gestual de señalamiento como indicación sistemática de tiempo, mientras que no existe en dicha lengua ningún recurso verbal específico alternativo que desempeñe la misma función semántica y gramatical. De ese modo, el gesto se convencionaliza y se integra a la gramática de la lengua al combinarse con palabras dentro de ciertas estructuras gramaticales, con el objeto de cumplir la función de referir a un momento del día o un transcurso de tiempo. Ello se logra combinando dos tipos semióticos diferentes: gestos y palabras, los que a su vez son vehiculizados en dos modalidades diferentes: la visual y la auditiva. El caso examinado obliga a reconsiderar la caracterización estándar del lenguaje verbal humano, redescribiéndolo, en el mismo sentido propuesto que en el trabajo de Scotto, como un sistema semiótico heterogéneo y multimodal. Sobre la base de la investigación realizada por Floyd, Audisio analiza el proceso de gramaticalización

del gesto deíctico, porque no solo muestra cómo se puede producir la integración de distintas modalidades en los sistemas lingüísticos, sino porque, incluso, permitiría dar cuenta de cómo el lenguaje verbal, como parte del proceso *online* de comunicación, integra en su estructura gramatical y léxica, de manera no redundante, vehículos propios de la modalidad gestual.

En “Esquematicidad mimética en las metáforas multimodales: contribuciones a la teoría de la metáfora conceptual”, Carolina Mahler examina las ventajas de una propuesta específica acerca de las *metáforas multimodales*, en el marco de la *teoría de la metáfora conceptual* (TMC), elaborada por Lakoff y Johnson a partir de los años 80 del siglo pasado. Dicha propuesta indaga sobre la estructura de las metáforas conceptuales multimodales, vinculando la corporeidad y la multimodalidad en la experiencia, para dar cuenta de una estructura cognitiva muy básica, habitualmente denominada “dominio-fuente”, que se proyecta sobre otra, conocida como “dominio-destino”. Desde una perspectiva diacrónica, la autora defiende la idea de que, como dominio-fuente, los “esquemas miméticos”, y no los “esquemas de imagen”, pueden explicar apropiadamente la estructuración de las metáforas multimodales, dado que el pensamiento esquemático mimético posee un vínculo más directo con la diversidad de la experiencia sensorio-motriz y son cognitivamente más básicos en el desarrollo ontogenético. Esta propuesta fortalece, según muestra Mahler, el enfoque corporeizado que aporta la TMC para entender ciertas formas básicas de cognición humana, pero, además, podría reflejar mejor la presencia de formas de cognición multimodal en la temprana ontogenia, un fenómeno del que también dan cuenta las investigaciones en psicología del desarrollo y neurociencias cognitivas, así como los enfoques corporizados de la cognición, más allá de la

TMC. De este modo, la propuesta amplía el alcance de las metáforas conceptuales desde el enfoque monomodal en el que recibieron su formulación clásica, hacia una perspectiva multimodal más amplia.

Laura Danón se ocupa de un tópico común a los estudios filosóficos y empíricos sobre el comportamiento humano no verbal y el comportamiento animal. En “Comportamientos expresivos naturales y contenidos mentales no proposicionales”, la autora se apoya en distintos trabajos de la filósofa Bar-On y colegas, quienes presentan un modo de interacción que puede tener lugar en criaturas carentes tanto de lenguaje como de pensamiento proposicional (en sentido pleno): la llamada “triangulación intermedia”. Quienes participan de este tipo de vínculo triangular con otro semejante y con un objeto del entorno común dan respuestas afectivas y comunicativas consistentes en la producción de *comportamientos expresivos naturales*. Tales comportamientos serían naturalmente diseñados para mostrar los estados psicológicos de quien los realiza a otros que estén dotados cognitivamente para reconocerlos. Bajo esta caracterización, los estados mentales que estas criaturas manifiestan mediante sus comportamientos expresivos son caracterizados por poseer contenidos no proposicionales. Danón evalúa de qué modo los autores mencionados consideran los contenidos no proposicionales que una criatura manifiesta mediante sus comportamientos expresivos y argumenta que son dos los posibles modos distintos de entender tal noción: como *contenidos de objeto* o como *contenidos que representan estados de cosas*, aunque ninguno de los dos llega a satisfacer todos los requerimientos de los contenidos proposicionales en sentido pleno. Planteada esta disyuntiva, examina cuál posición es más adecuada, y analiza, para ello, tres vías alternativas que estos autores podrían seguir: a)

entender la totalidad de los contenidos que manifiestan los comportamientos expresivos como contenidos de objeto; b) entenderlos como contenidos que representan estados de cosas, pero no llegan a ser contenidos proposicionales plenos; o bien, c) adoptar un enfoque híbrido, de acuerdo con el cual los comportamientos expresivos pueden manifestar, en distintas ocasiones, tanto estados mentales con contenidos de objeto, como contenidos no proposicionales (en el sentido pleno del término) pero que aún versen sobre estados de cosas. Según sugiere la autora, esta última noción es la más adecuada si se quiere capturar el amplio espectro de conductas y capacidades cognitivas presentes en distintas criaturas sin lenguaje.

En “Algunas reflexiones en torno a las otras mentes”, Daniel Kalpokas presenta algunas observaciones acerca del problema del conocimiento de las otras mentes. Tales observaciones, aclara Kalpokas, se sitúan en el nivel personal de análisis, y conciernen al problema epistemológico de cómo, a partir de las conductas expresivas de los otros, logramos conocer sus estados mentales. Después de presentar un argumento, basado en la fenomenología de la experiencia perceptiva, a favor de la tesis de que, al menos en ocasiones, podemos percibir los estados mentales ajenos, Kalpokas hace una propuesta acerca de cómo podría ser entendida dicha tesis. Esta propuesta se articula sobre la base de dos tesis: una acerca de las conductas expresivas de los agentes, y otra acerca de la percepción de los estados mentales ajenos. En primer lugar, sugiere que entendamos las conductas expresivas que los otros realizan sin intención comunicativa como *que significan naturalmente* sus estados mentales correspondientes. Puesto que, según se suele entender, el significado natural es factivo, si x (cierta conducta) significa naturalmente p (cierto estado mental), entonces x implica p . Así, si una persona se comporta de

un modo que significa naturalmente cierto estado mental, entonces podemos estar seguros de que tal persona está en tal estado mental. La idea principal es que percibir ciertos comportamientos como que expresan tal y cual estado mental es percibirlos como que significan naturalmente tal y cual estado mental. El conocimiento de trasfondo de que tales y cuales comportamientos significan naturalmente tales y cuales estados mentales es lo que, penetración cognitiva mediante, explica la fenomenología de la experiencia cuando experimentamos las conductas de los otros en cuanto expresan dichos estados. La segunda tesis, que retoma una línea conocida en filosofía de la percepción, sostiene que esta es una función de las expectativas que típicamente nos formamos cuando percibimos un objeto desde cierta perspectiva y en ciertas condiciones ambientales. Dichas expectativas conciernen a cómo esperamos que se nos aparezca un objeto en caso de que cambiemos de posición respecto a él. En la percepción encontramos tanto lo directamente presentado (la cara del frente de un objeto, por ejemplo) como lo copresentado (la cara posterior del objeto). En el caso de las conductas expresivas, los comportamientos directamente percibidos copresentan otros comportamientos que son constitutivos, en cada caso, de los tipos de eventos que son. El capítulo cierra respondiendo algunas posibles objeciones a esta propuesta.

En “Percepción dinámica de expresiones faciales y atribución de emociones”, Zoé Sánchez Barbieri argumenta a favor de considerar la dinámica del comportamiento expresivo como un factor esencial para explicar la habilidad para reconocer emociones. La autora examina cómo la detección por parte del intérprete del despliegue secuencial de expresiones faciales y corporales de un individuo le permite identificar de manera apropiada el *tipo* de estado emocional que experimenta el individuo y posibilita,

más básicamente, que el intérprete pueda distinguir las emociones de otro tipo de estados psicológicos. A fin de alcanzar el objetivo propuesto, la autora sigue los siguientes pasos en su examen: en primer lugar, caracteriza brevemente algunas metodologías usuales que los psicólogos cognitivos han utilizado para evaluar el reconocimiento de emociones, basadas en imágenes estáticas captadas mediante fotografías de distintas expresiones faciales características. Luego, identifica las dificultades más importantes que derivan de los protocolos realizados con este tipo de estudios con imágenes, principalmente aquellas basadas en la evaluación desprovista de contexto de las expresiones faciales y corporales. Sánchez Barbieri señala la importancia de tomar en cuenta distintos elementos contextuales como elementos fundamentales para explicar nuestra habilidad para reconocer emociones. Por último, examina la dinámica de las expresiones faciales y la distinción proceso-producto aplicada a estos fenómenos expresivos, subrayando cómo el registro del proceso emocional, por parte del intérprete, tiene una importancia fundamental para dar cuenta del reconocimiento de emociones en nuestras interacciones cotidianas. La combinación de los rasgos dinámicos y los factores contextuales, los que a su vez ponen de relieve el papel de los contextos interactivos en los que tiene lugar –regularmente– el reconocimiento de emociones, abona una perspectiva interaccionista de la atribución intencional, al menos para los casos más básicos.

El conjunto restante de los trabajos recoge resultados de sendas investigaciones empíricas relacionadas con las interacciones tempranas en general, el reconocimiento de emociones, la recíproca influencia entre la percepción de propiedades amodales y formas de vínculo adulto-bebé en el desarrollo de las habilidades semióticas. Algunos de

los textos se concentran en la discusión teórica de tópicos como el innatismo o el modularismo de ciertas competencias psicológicas de la primera infancia, otros apuntan a un enlace con el campo clínico y el pedagógico.

En “Desarrollo semiótico, multimodalidad y cuerpo”, Fernando G. Rodríguez se propone, recapitulando algunos puntos ciegos de la versión modular-innatista del lenguaje, revisar la serie de los hitos del desarrollo sociocognitivo de la primera infancia, particularmente aquellos que atañen a la ontogénesis de las habilidades expresivo-comunicativas, considerando en detalle la condición del gesto como primer signo de realización deliberada en el sujeto humano y su ulterior (en realidad, casi contemporánea) relación de acoplamiento con la palabra. Los progresos de la semiosis infantil son presentados desde una perspectiva que entiende de las competencias de interacción en general, tanto las asociadas con la intersubjetividad primaria como las que normalmente se han reconocido decisivas de la secundaria, no solo en cuanto precursoras, sino en calidad de factores determinantes en el ingreso del niño al mundo de la significación intencionada. Esta es asumida como el desenlace de la conjunción de muy distintas destrezas antecedentes, de modo especial las vinculadas al uso del cuerpo y a la gesticulación. De modo particular, el texto se concentra en las contribuciones específicas de los gestos de tipo representacional al desarrollo comunicativo del sujeto. Aunque este gesto constituye una categoría habitual en la investigación empírica, no es por lo regular considerado en el nivel teórico, desde una perspectiva de conjunto donde las distintas estrategias y recursos comunicativos del niño pequeño constituyen *de hecho* un elenco integral multivariado. Precisamente, dentro de la cartografía de la semiosis del sujeto, los gestos representacionales se dejan reconocer como signos irreductibles, en su forma y

su materia, a otros tipos de gestos y a la verbalización. En la interfaz entre el mundo referencial (esto es, el mundo posible de ser significado) y el cuerpo como instrumento, capaz de escenificar o modelar la realidad de los objetos, los gestos representacionales implican un recorte o selección del cuadro perceptivo, y luego una expresión, exteriorización isomórficamente elaborada. El gesto de esta clase supone, por tanto, la variante sínica en que la plasticidad mimética del cuerpo alcanzaría el punto más alto de iconicidad multimodal que puede registrarse en los actos de significación humana. El recorrido pretende acentuar la máxima importancia que reviste la reintroducción de la corporeidad en los estudios sobre el desarrollo semiótico en sentido lato, y de un modo concreto en los estudios sobre las primeras fases del lenguaje concebido como actividad (*enérgeia*).

En “Desarrollo intersubjetivo y perceptivo: el ritmo como ejemplo de su enlace durante el primer año de vida”, Mauricio Martínez discute el tránsito desde la intersubjetividad primaria a la secundaria (revolución del noveno mes) señalando que la mayoría de las teorías que abordan este pasaje se excusan con elegancia de dar una explicación de él, e invariablemente aluden a cierto salto cualitativo. Advierte, sin embargo, que a lo largo del primer año de vida el niño experimenta, sobre todo, cambios paulatinos en los modos de su interacción. Según su tesis:

Si asumimos que existe un vínculo entre la percepción intersensorial y el desarrollo intersubjetivo, resultaría plausible interpretar que los cambios en la conducta del bebé se encuentren, de algún modo, sustentados por cambios en sus capacidades perceptivas. Ahora bien, [...] ¿qué clase de información percibe el bebé durante los intercambios?, [y] ¿qué clase de relación debemos establecer entre el desarrollo de ambas capacidades?

La clave propuesta para cruzar ambas competencias, una cognición *fría* y otra templada en el calor del intercambio socioemocional, reside en una de las facetas de la musicalidad comunicativa, a saber, el ritmo, donde el autor cree hallar el hilo conductor que atraviesa una serie de modificaciones en las pautas perceptivo-relacionales del bebé. Mes a mes puede rastrearse la evolución de la conducta infantil frente a los modelados rítmicos multimodales del adulto. Dicho con una fórmula: el involucramiento del bebé en la interacción diádica primaria permitiría el desarrollo de su percepción intersensorial y, dialécticamente, esta iría modelando las formas de su socialidad en un marco ascendente, espiralado, de recíproco apuntalamiento. La indicación es de primera relevancia, ya que pretende llenar el espacio entre intersubjetividad primaria y secundaria desde una psicología del desarrollo de estricto sentido empírico, sin echar mano al recurso habitual de suponer, a esta temprana edad, dispositivos de carácter dado y ya acabados al momento de nacer.

En “Competencia social y reconocimiento de emociones a partir de expresiones faciales y lenguaje corporal en sujetos del espectro autista. Plataformas virtuales como apoyo a la experiencia”, Xilenia Carreras examina el reconocimiento, la transmisión y la comunicación de estados emocionales a partir del lenguaje corporal y la “lectura” de las expresiones faciales en sujetos afectados por la condición del espectro autista (CEA). Sobre la base de un modelo corporizado de la cognición social, que busca explicar de qué modo los niños con desarrollo típico consiguen reconocer emociones y comunicarse eficazmente con otros, la autora analiza algunos de los déficits en la interacción social que caracterizan a los sujetos con CEA, que consisten en alteraciones clínicamente significativas en la interacción social y la comunicación, y que

se manifiestan mediante patrones restrictivos o repetitivos del comportamiento, los intereses o las actividades. A partir del marco teórico escogido, la autora se propone evaluar algunos de los usos, ventajas y limitaciones derivados del empleo de plataformas virtuales para la enseñanza y el entrenamiento de habilidades socioemocionales, considerando experiencias y desarrollos específicos. Como consecuencia de ello, finalmente, realiza algunas sugerencias para el diseño y el empleo de plataformas virtuales que pueden servir de soporte para enriquecer la experiencia social en estos casos, tanto en contextos clínicos como educativos.

El trabajo “Gestos y palabras. Los gestos de maestras y alumnos durante la enseñanza de vocabulario” de Alejandra Menti está dedicado al análisis de la función que cumplen los gestos de docentes y alumnos durante el aprendizaje de vocabulario específico, no familiar o poco frecuente, en clases de Ciencias Sociales en la escuela primaria. Se trata de una investigación empírica, realizada en el ámbito de las escuelas públicas primarias de Córdoba (Argentina), para la cual se adoptó una metodología mixta que combina de modo secuencial una fase cualitativa y otra cuantitativa. Los resultados del análisis cualitativo pusieron de manifiesto que tanto las maestras como los niños emiten gestos con cuatro distintas funciones: ampliar o reforzar la información que expresan verbalmente, cubrir el vacío lingüístico, manifestar comprensión compartida y enfatizar aspectos de sus intervenciones. Los alumnos se valen especialmente de gestos icónicos ante dificultades en la recuperación del léxico técnico apropiado, mientras que los docentes utilizan estos mismos gestos para favorecer fáticamente la interacción y consolidar el concepto en el niño, si su gesto hubiera sido conceptualmente atinado. El análisis cuantitativo mostró que los

interlocutores –docentes y alumnos– privilegian el empleo de gestos destinados a ampliar o reforzar la información que proporcionan verbalmente. La investigación pone de manifiesto que el aprendizaje de vocabulario conceptual no depende exclusivamente del discurso verbal, de la cantidad, diversidad o complejidad de los términos del léxico, sino del aporte en simultáneo de la gesticulación, sugiriendo que, como en otros campos, también en el espacio de la educación el gesto cumple un rol central, permitiendo, a los docentes, optimizar la transmisión de contenidos y, a los alumnos, consolidar el proceso de incorporación de conceptos. Este texto, ya publicado previamente, constituye el informe final de una investigación concluida y un valioso reflejo de cómo se estudia empíricamente la presencia del gesto en un espacio práctico, en este caso la escuela, habilitado entonces como herramienta fecunda para repensar las coordenadas básicas de la transmisión de conocimientos.

Los trabajos reunidos se ubican en una posición que comprende la mente de forma no restringida a la computación de símbolos abstractos. Al contrario, desde el trabajo empírico y desde la evaluación teórica, defienden por igual la primera necesidad de entender que otros aspectos deben integrarse a la investigación sobre la mente y la comprensión de otras mentes, superando la oposición yo-otro, desentendida de los puntos de encuentro *inmediato* que enlazan las subjetividades. Estos puntos de encuentro *precognitivos* (si reducimos falsamente los procesos cognitivos a procedimientos de tipo informático, según la metáfora del pensamiento como un algoritmo de suprema densidad) tienen su espacio en la corporeidad, plasmada en el rostro, en la expresión facial, en manifestaciones

espontáneas o deliberadas de emociones, en la gestualidad, en las habilidades perceptivas. De ese modo, tendríamos acceso a un otro que no sería todavía, en los comienzos de la interacción diádica con semejantes, un otro basado en atribuciones mentalistas, con una mente *interior*, sino una alteridad *plena* y sin doble fondo, sin desdoblamiento entre fenomenidad patente y un lado de atrás solo indirectamente penetrable. La reivindicación del cuerpo como parte esencial de los problemas referidos a la condición humana en su nivel más básico parece, a la luz del que ha sido llamado “giro corporal”, una demanda generalizada ante el agotamiento de una perspectiva para la que no hay que confundir limitaciones con falta de méritos. La cuantiosa contribución del cognitivismo mentalista, por espacio de más de medio siglo, ha dado así su mejor fruto desde el ángulo particular en que se había ubicado, y ahora se mueve, transformado, hacia una posición desde la que el objeto le revela aristas previamente ocultas. Pero inclusive el lado oculto de la Luna se descubre en cuanto nos desplazamos, y es justamente de ello, de movimiento, de lo que se trata. De cómo variables como el movimiento, que es movimiento corporal en el espacio y en el tiempo, condicionan nuestra percepción y posterior conocimiento del objeto.

Parte I. Multimodalidad y lenguaje

El lenguaje verbal también es icónico: correspondencias transmodales y simbolismo sonoro

CAROLINA SCOTTO (UNC – CONICET)

1. Iconicidad vs. arbitrariedad en el lenguaje

La existencia de relaciones icónicas o de semejanza entre los signos o vehículos lingüísticos y los significados se puede rastrear en una antigua tradición cuyo primer expositor notable fue Platón. En el debate entre naturalismo y convencionalismo¹, Crátilo defiende, frente al sofista Hermógenes, la tesis según la cual los nombres significan lo nombrado por medio de una relación “natural” imitativa (*Crátilo*, 423b), que se extiende a los elementos subléxicos que los componen: fonemas, vocales y consonantes, así como a los rasgos prosódicos y a la métrica. Esta idea fue ricamente elaborada en la teoría semiótica de Peirce, a comienzos del siglo XX, y retomada en la lingüística, entre otros, por Jespersen (1921) y más tarde por Jakobson (1965). Sin embargo, en los estudios filosóficos y lingüísticos se consolidó una tradición diferente, que alcanzó su

¹ Desde el *Crátilo*, es frecuente confundir convencionalidad con arbitrariedad: si las palabras significan por convención, entonces parece natural suponer que son arbitrariamente elegidas. Sin embargo, las palabras convencionales de una lengua pueden tener un origen icónico (Ahler y Zlatev, 2010).

formulación ortodoxa en la obra fundacional de la lingüística contemporánea, el *Cours de Linguistique Générale* de F. de Saussure (1916)². En esta otra tradición, se afirma, por el contrario, que la mayor parte de los signos lingüísticos se vinculan con sus significados por medio de una relación arbitraria y que, como consecuencia, los fenómenos icónicos solo pueden tener una significación secundaria o marginal en el lenguaje. Esta es, claramente, la “visión recibida” (Perniss, Thompson y Vigliocco, 2010).

En este trabajo me propongo desarrollar un argumento empírico, que tiene como premisa la abundante evidencia que corrobora la *ubicuidad* y la *importancia relativa* del simbolismo sonoro en el lenguaje, entendido como un fenómeno icónico, para defender una caracterización radicalmente diferente del lenguaje natural de la que ofrecen las teorías o enfoques “ortodoxos”, tanto en la lingüística como en la filosofía del lenguaje. Dicho argumento tiene la siguiente forma: si el simbolismo sonoro, como un subtipo lingüístico de iconicidad, es una *propiedad general del lenguaje* y, además, es una expresión del papel de las correspondencias transmodales en el lenguaje, entonces el lenguaje debería ser redescrito como un *fenómeno semiótico heterogéneo y multimodal*. En segundo lugar, sobre la base de esta conclusión, dado que en el conocimiento y uso del lenguaje están involucradas dimensiones icónicas y multimodales, estas serán también relevantes para explicar una variedad de otros aspectos, tales como el aprendizaje y el procesamiento lingüístico, los usos comunicativos, los orígenes filogenéticos, el desarrollo ontogenético, el cambio lingüístico, entre otros. No es objetivo de este trabajo identificar y justificar los efectos que podrían derivarse

² Locke (1690) proporcionó un argumento muy persuasivo en su favor: si los nombres estuvieran conectados naturalmente con sus significados, todos hablaríamos la misma lengua, lo que es manifiestamente falso (Cfr. Perniss *et al.*, 2010).

de la reconfiguración global propuesta, la cual, además, depende de la consideración de otros rasgos que también forman parte del paradigma “ortodoxo”³. Solo intentaré apoyar esa reconfiguración en la valoración de algunas importantes comprobaciones empíricas acerca de cuál podría ser el papel del simbolismo sonoro en el aprendizaje temprano del lenguaje. Cabe agregar, finalmente, que el análisis del simbolismo sonoro será inscripto en el marco más amplio del fenómeno conocido como “correspondencias transmodales” (CT), a cuyas características me referiré en primer lugar. Como consecuencia de ello, nuestro tema quedará enmarcado en los vínculos entre percepción (o sistema sensorio-motor) y lenguaje, en un claro contraste con las perspectivas tradicionales, que vinculan al lenguaje solo o primariamente con la cognición. En este punto tiene su origen el cambio de enfoque propuesto: puesto que los fenómenos sonoro-simbólicos son, en parte, fenómenos de correspondencias entre modalidades, reconocer la importancia de estas para la cognición humana en general permitirá comprender que el simbolismo sonoro, lejos de tener una importancia marginal, es, por el contrario, una manifestación en el lenguaje de capacidades cognitivas generales, más específicamente perceptivas, que exceden al lenguaje. Como veremos también, la naturaleza de estas CT obliga, además, a adoptar un enfoque no tradicional de la percepción sensorial.

³ Nos referiremos aquí sobre todo a la arbitrariedad y, en menor medida, al canal vocal-auditivo. Sin embargo, de acuerdo con los enfoques ortodoxos, también otros rasgos serían distintivos de los lenguajes humanos. Un ejemplo de estos enfoques se encuentra en Hockett (1960).

2. ¿Qué son las correspondencias transmodales?

Se denominan “correspondencias transmodales” (CT)⁴ a una variedad de percepciones multisensoriales en las cuales estímulos o atributos de esos estímulos, no necesariamente presentes y aparentemente no relacionados entre sí, en diferentes modalidades y dimensiones sensoriales, producen una concordancia, conjunción o “efecto de congruencia” o de compatibilidad, generalmente bidireccional (Deroy y Spence, 2016), en el contenido de la experiencia, denominada *congruencia sinestésica* o *congruencia semántica* (Spence, 2011; Sidhu y Pexman, 2017b). De ese modo, dichas correspondencias resultan ser consistentes o “esperadas” y no arbitrarias (Parise y Spence, 2012). Por ejemplo, se asocian consistentemente sonidos de tono agudo con objetos pequeños, o sonidos fuertes con formas u objetos grandes (esta última, constatada incluso en niños de 2 años de edad) (Spence, 2011); también suele producirse una asociación entre tonos altos y luminosidad o brillo (incluso en bebés muy pequeños) (Walker *et al.*, 2010); entre timbres musicales y olores; formas y colores; colores y temperatura, música, olores, formas, sabores y letras, etc. En suma, las asociaciones comprenden desde estímulos muy elementales (o de “bajo nivel”), como la luminosidad, hasta otros muy complejos (o de “alto nivel”), como pseudopalabras (Parise, 2016). Una variedad de tests experimentales, basados en la velocidad para asociar o clasificar, la automaticidad y la recurrencia de las respuestas para seleccionar y asociar estímulos (Parise y Spence, 2012), convergen en señalar la robustez del fenómeno,

⁴ En la literatura se encuentran muy distintas denominaciones del mismo fenómeno, por ejemplo: “congruencias sinestésicas”, “asociaciones sinestésicas”, “equivalencias transmodales”, “similitudes transmodales”, “mapeos transmodales naturales”, entre otras (ver Spence, 2011, pp. 972-973).

así como también su presencia generalizada en todas las poblaciones estudiadas, es decir, en distintos grupos etarios y comunidades lingüísticas y culturales. Dicha evidencia desmiente la presunción de que podría tratarse o bien de un fenómeno limitado a la población de sinestésicos⁵ (Parise y Spence, 2012), o bien a experiencias anecdóticas, idiosincráticas o solo culturalmente acotadas, de alcance e interés teórico más limitado (Deroy y Spence, 2016).

El fenómeno fue identificado a comienzos del siglo pasado por Köhler (1929), pero es objeto de un interés teórico creciente, en disciplinas como las neurociencias, la psicología cognitiva, la psicología del desarrollo y la lingüística, desde que la publicación de Ramachandran y Hubbard (2001) puso el tema nuevamente en el centro de la escena. Sin embargo, su antecedente más claro se relaciona con los estudios acerca del problema de la integración multisensorial, el así llamado “problema del *binding*”, que consiste en explicar cómo podemos percibir objetos, segregando e integrando propiedades, a pesar de que procesamos separadamente, en distintos circuitos neuronales, un *input* sensorial complejo, es decir, una información sensorial variada (colores, tamaños, formas, etc.). Sin embargo, el problema de la integración multisensorial, al ocurrir para una clase más amplia de estímulos perceptuales, es decir, al no ser solo acerca de propiedades de un mismo objeto, sino también acerca de propiedades que

⁵ Son sinestésicos los individuos (según algunas estimaciones, aproximadamente el 4 % de la población) que experimentan vívidamente perceptos multimodales, v.gr. saborean colores, ven sonidos, etc. A diferencia de las experiencias “normales”, las sinestésicas agregan a la experiencia relacionada con el estímulo que la provoca, un percepto adicional o “concurrente”, a veces en una diferente modalidad sensorial. Es interesante que el mayor volumen de experiencias sinestésicas estén desencadenadas por estímulos lingüísticos (Simner, 2007).

son percibidas como pertenecientes a objetos distintos, exceden al “problema del *binding*” (Deroy y Spence, 2016; Spence, 2011).

Si las CT son comunes a toda la población y son características de nuestra actividad perceptiva habitual, entonces la *concepción unimodal* clásica de la percepción (O’Callaghan, 2012) debería considerarse, cuanto menos, parcialmente errónea. Por lo tanto, investigar la percepción separadamente en cada una de las modalidades, asumiendo que pueden explicarse con independencia de sus vínculos con las otras, y presuponer, además, que la sumatoria de las historias específicas acerca de los rasgos de cada modalidad sensorial dará por resultado una teorización exhaustiva acerca de la percepción (O’Callaghan, 2012) no es compatible con diversos fenómenos revelados por las CT. La atención monopólica prestada a la visión por sobre otras modalidades sensoriales (el llamado “visuocentrismo”) y la proyección injustificada de sus rasgos específicos a las demás modalidades profundizaron el carácter distorsivo del enfoque clásico de la percepción. La concepción unimodal parece respaldada intuitivamente por el hecho de que solo podemos comparar apropiadamente los estímulos si pertenecen a una misma modalidad, en tanto ocasionan experiencias cualitativamente semejantes entre sí (*qualia unimodales*). Las teorías empiristas clásicas de la percepción sostuvieron, sobre estas bases, que la percepción de rasgos atómicos (como el color, la forma, etc.) requería ser complementada por procesos asociativos u otros en la mente, responsables de reunir esa información en totalidades integradas. Este enfoque contrasta con otro más reciente, según el cual la percepción es, en general, “ricamente multimodal”, en todos los niveles explicativos, incluido el fenomenológico (O’Callaghan, 2012). Dado que normalmente percibimos con muchos

sentidos, y que, además, estos se activan regularmente en contextos multimodales, lo que percibimos es, simplificadamente, un efecto natural de la cooperación entre distintas modalidades sensoriales. Esta produce experiencias o bien *integradas*, cuando son acerca de un mismo objeto, o bien *unificadas*, cuando son acerca de distintas entidades que interactúan.

En la literatura se ha propuesto distinguir entre dos tipos básicos de CT. Por una parte, la denominada “congruencia sinestésica”, que serían las correspondencias no arbitrarias entre los rasgos o atributos físicos más básicos de los estímulos, tanto estímulos no redundantes (aque-lllos que solo pueden ser percibidos por una modalidad, por ejemplo, el brillo de un objeto y el tono de un sonido), como redundantes (rasgos que podrían ser percibidos a través de distintas modalidades, por ejemplo, el tamaño o la forma de un objeto percibidos por la vista y por el tacto, o la duración visual y auditiva de un evento). Estas congruencias no solo ocurren en personas sinestésicas, sino que muchas de ellas son comunes, en diversos grados, a la mayoría de las personas. Se ha propuesto la noción de *sinestesia débil* para cubrir este tipo de correspondencias, la que integraría un *continuum* gradual con la *sinestesia fuerte*. Ambas refieren a las mismas correspondencias, aunque con una “fuerza” diferente (Martino y Marks, 2001)⁶. Aunque no todas las correspondencias son universales, algunas quizás sí lo sean, como parecería evidenciarlo su detección en niños muy pequeños. Por otra parte, la llamada “congruencia semántica” refiere a estímulos de diferentes modalidades (típicamente visual y

⁶ Aunque otros autores cuestionan que la sinestesia sea un buen modelo para pensar las correspondencias transmodales, porque en aquella se produce la correspondencia de manera unidireccional, lo que genera una experiencia “concurrente” (Parise y Spence, 2012, p. 330).

auditiva) que varían de maneras congruentes relativamente a su identidad o pertenencia a una u otra categoría, o a su significado (por ejemplo, imágenes y sonidos, letras y sonidos de habla, género y palabras). Otros autores la denominan “sinestesia lingüística” o “conceptual” (Simner, 2007). Las CT del primer tipo no están mediadas por conceptos, y podrían ser más universales que aquellas que sí lo están.

Respecto de los mecanismos que podrían explicar estas CT, como dijimos, las primeras hipótesis apelaban a la congruencia espacio-temporal (estímulos simultáneos y coincidentes espacio-temporalmente), pero también se han estudiado otros factores y mecanismos causales. La hipótesis de la *correspondencia estadística* apela a una co-ocurrencia de los estímulos en la naturaleza, esto es, el vínculo estable entre las propiedades de los objetos en sus escenarios habituales (por ejemplo, los objetos grandes en tamaño o masa producen regularmente sonidos más fuertes en volumen). Otra hipótesis busca los fundamentos de estas correlaciones no en la naturaleza, sino en las conexiones neurales innatas en el sistema perceptual, denominada “correspondencia estructural” (Emmory, 2014). Estos dos tipos de correspondencias podrían, a su vez, explicarse de manera complementaria (Parise, 2016). Finalmente, se ha propuesto explicar las correlaciones por medio de las palabras usadas regularmente para vincular los estímulos: por ejemplo, “alto-bajo”, aplicado al tono de un sonido y a las dimensiones de un objeto. Se la denomina “correspondencia semánticamente mediada”⁷. Un subtipo de estas podría estar constituido por las

⁷ Es interesante esclarecer la relación entre estas correspondencias semánticamente mediadas con las metáforas cognitivas, que operarían también mediante cruces entre modalidades sensoriales, moduladas por nuestra experiencia sensorial y corporizada (Forceville, 2006).

“correspondencias hedónicamente mediadas”, que consisten en asociaciones entre dos estímulos mediados por un estado emocional: por ejemplo, entre música y colores, o rostros expresivos y colores o música (Palmer, Schloss, Xu y Prado-León, 2013). Como veremos, mecanismos similares se postulan para explicar el simbolismo sonoro, lo que no debería llamar la atención si se entiende que este es una manifestación particular de las CT, aun cuando, además, excede el dominio de los fenómenos abarcados por estas. De todas ellas, las correspondencias estructurales son las más probablemente innatas, o menos susceptibles de ser modificadas por el aprendizaje y la cultura, a diferencia de las restantes⁸. No obstante, aún no hay suficientes investigaciones sobre el papel de la experiencia, los cambios en el desarrollo y el aprendizaje que permitan calibrar en qué medida estos factores pueden explicar o bien modular las CT. Del mismo modo, aún restan por explorarse una multiplicidad de asociaciones respecto de diversos parámetros. En síntesis, los resultados de los experimentos realizados hasta ahora admiten una variedad de criterios para clasificar las correspondencias y una variedad de mecanismos que podrían complementarse, así como diversos problemas o desafíos aún pendientes por resolver acerca de ellas.

⁸ Se han constatado tanto semejanzas como diferencias culturales en las asociaciones entre sabores y otras dimensiones no gustativas, como colores y formas, comparando poblaciones de China, India, Malasia y EE.UU., con muy diferentes patrones de cultura alimentaria (Wan *et al.*, 2014).

3. ¿Qué es el simbolismo sonoro?

Desde sus inicios, se advirtió que también las palabras, habladas o escritas⁹, se cuentan entre los estímulos que provocan CT, en cuanto son objetos de percepción, auditiva o visual (Sapir, 1929; Köhler, 1929)¹⁰. El fenómeno lingüístico se denominó “simbolismo sonoro” (o “simbolismo fonético”)¹¹ (Jespersen, 1921), y su naturaleza y variedades son todavía materia de investigaciones e hipótesis alternativas. La conjetura, originalmente formulada por Jespersen (1921), se basó en el descubrimiento de que el sonido vocálico /i/ se asocia regularmente con las cosas pequeñas, y los sonidos vocálicos /u/, /o/, /a/, con las grandes, y que estas asociaciones podrían deberse a la forma como se disponen la cavidad bucal y la lengua para su pronunciación. Se afirma que estas correspondencias son universales. Muchas otras correspondencias han sido descriptas con posterioridad, abarcando una amplia variedad de dimensiones y dominios semánticos, tales como líquido-sólido, delgado-ancho, rápido-lento, abrupto-continuo, alto-bajo, activo-pasivo, etc. (Perlman,

⁹ En adelante nos referiremos solo a la modalidad vocal del lenguaje, por tratarse de la manifestación primaria del lenguaje verbal, pero la iconicidad ha sido también muy estudiada en las formas gráficas de la escritura, es decir, en la modalidad visual.

¹⁰ En la literatura todavía no existe consenso acerca de cuáles fenómenos son abarcados por las categorías de CT y simbolismo sonoro. En parte por esa razón, la inclusión del simbolismo sonoro dentro de las correspondencias transmodales tampoco está del todo consolidada (cfr. Sidhu y Pexman, 2017b).

¹¹ Muchos autores han señalado la inconveniencia de mantener estas denominaciones históricas, en tanto se desea preservar la estrecha relación de este fenómeno con las ideas peirceanas sobre la iconicidad. Como sabemos, Peirce define a los *símbolos* por contraste con los *íconos*, como aquellos signos que no guardan ninguna relación de semejanza con aquello que significan (cfr. Ahler y Zlatev, 2010, p. 312). En cualquier caso, las denominaciones en la literatura son muy variadas y discutidas: “iconicidad sonora”, “iconicidad sonoro-sinestésica”, “iconicidad fonémica”, “asociación sonido-significado”, “iconicidad transmodal”, entre otras.

2017). El simbolismo sonoro ha sido ampliamente constatado en diferentes rasgos de la modalidad vocal-auditiva, tanto prosódicos como articulatorios, a nivel submorfémico y morfémico, así como también a nivel de las palabras y de construcciones lingüísticas complejas, e incluso en diferentes dimensiones: sintáctica, semántica y pragmática (Dingemanse, Blasi, Lupyan, Christiansen y Monaghan, 2015). En términos muy generales, su caracterización en la literatura oscila entre una versión más débil y otra más fuerte: o bien serían *asociaciones regulares* (o “co-ocurrencias estadísticas”) entre formas lingüísticas (sonoras o gráficas¹²) y significados, o bien se trataría de *relaciones icónicas* más robustas, por las cuales los signos o formas acústico-articulatorias reflejan o “imitan” algún aspecto o propiedad del referente, del significado o de las experiencias de sus usuarios ligadas consistentemente a su empleo (Hinton, Nichols y Ohala, 1994). Ahora bien, la mayor parte de los autores coincide en que el fenómeno manifiesta variaciones graduales en un continuo en el que intervienen elementos más objetivos o más subjetivos, y factores más innatos o más aprendidos, entre otros, así como también en que atañe a una variedad de propiedades e involucra tipos de semejanzas también diferentes. Estas correspondencias, a nivel de los signos en general, lingüísticos o no, han sido caracterizadas como icónicas. De acuerdo con Peirce, un signo es icónico si está motivado, directa o indirectamente, por la semejanza con aquello que representa. En ese sentido, está justificado subsumir bajo el concepto de “iconicidad” a las diversas manifestaciones de las relaciones de semejanza entre signos de cualquier modalidad lingüística y sus significados (Perniss y Vigliocco, 2014). En síntesis, entendemos por iconicidad

¹² La mayor parte de los autores recientes incluyen también a los signos icónicos de los lenguajes de señas. Volveremos luego sobre este tema.

a la propiedad de ciertos vehículos o de la forma lingüística –o, más ampliamente, de la forma comunicativa (incluyendo a las señas de los lenguajes de señas y también a los gestos que acompañan al habla)– de poseer algún tipo y grado de semejanza con ciertas propiedades sensoriomotoras de los referentes, o con lo que significan, o con las experiencias afectivas u otras evocadas por ellos (Nöth, 2015; Perniss y Vigliocco, 2014; Winter, Perlman, Perry y Lupyan, 2017).

Hasta hace relativamente poco tiempo se creía que las palabras icónicas eran solo una pequeña proporción del vocabulario total de las lenguas¹³, “casi un rasgo paralingüístico” (Perniss *et al.*, 2010), y que, por esa razón, servían ante todo para confirmar la tesis acerca de los lazos arbitrarios entre signos y significados en la inmensa mayoría de las restantes palabras de los lenguajes naturales¹⁴. La arbitrariedad consiste en que “cualquier combinación de sonidos puede significar cualquier contenido semántico” (Dingemanse *et al.*, 2015, p. 603), y, en consecuencia, dado el sonido de una palabra desconocida, no es posible inferir su significado (Monaghan, Shillcock, Christiansen y Kirby, 2014). La arbitrariedad, de acuerdo con la formulación saussuriana, sería una característica definitoria, e incluso “un primer principio” del lenguaje humano. Otra fuerte referencia, también en el ámbito de la lingüística,

¹³ En realidad, es todavía opinión corriente u “ortodoxa” la omnipresencia de la arbitrariedad, tanto en la lingüística teórica, sobre todo en la vertiente chomskiana, como en la filosofía analítica del lenguaje, lo que se refleja en la ausencia del tópico en la reflexión filosófica sobre el lenguaje, después de Peirce.

¹⁴ Las relaciones forma-significado también pueden ser *sistemáticas* en los casos en que ciertos patrones sonoros, fonológicos y prosódicos se asocian regularmente a ciertas categorías sintácticas (verbos vs. nombres) o semánticas (familias de significados). No me ocuparé aquí de la sistematicidad, pero es importante advertir que, a diferencia de la arbitrariedad, los rasgos sistemáticos también proporcionan claves acerca del significado, ya que permiten derivar la categoría gramatical de la forma de la palabra y con ello facilitan la categorización y el aprendizaje (Dingemanse *et al.*, 2015).

se puede encontrar en la clásica obra de Hockett (1960, 1963)¹⁵, quien propuso un conjunto de “principios de diseño” o de rasgos universales que serían comunes a todos los lenguajes humanos. Para nuestros propósitos, es importante señalar dos de ellos, dado que, como veremos, están estrechamente vinculados: a) la utilización del canal vocal-auditivo; y b) el carácter arbitrario de los vínculos entre signos y significados.

Sin embargo, aunque para muchos la arbitrariedad es una propiedad evidente o intuitiva si se toma una variedad de palabras como ejemplos, esta ha sido gradual y parcialmente desmentida, en cuanto “principio” o rasgo definicional, por las evidencias convergentes surgidas del estudio de las más diversas lenguas de distintas familias, de todo tipo de palabras y en todas las dimensiones lingüísticas, facilitados por la disponibilidad de *corpora* lingüísticos en bases de datos de gran escala (Christiansen y Monaghan, 2015). Estas lenguas han sido estudiadas mediante una batería de tests realizados a distintas poblaciones de hablantes en relación con distinto tipo de palabras, de la propia lengua o de lenguas desconocidas, o incluso de pseudopalabras¹⁶, que arrojaron respuestas consistentes sobre el reconocimiento de sus significados. Estas evidencias no solo han apoyado la tesis de que la iconicidad tiene un papel más importante que el que se le atribuía, sino que, incluso, debería considerarse también una *propiedad general del lenguaje*, o que, junto a la

¹⁵ Los rasgos de diseño propuestos por Hockett variaron en número en distintas formulaciones (1958, 1960, 1962, 1963, 1968), pero la arbitrariedad, que es la que principalmente nos interesa aquí, permanece en todas ellas.

¹⁶ Los primeros experimentos de Sapir y Köhler, emplearon pseudo-palabras, esto es, palabras inventadas pero fonológicamente típicas de una lengua dada, utilizadas para evitar la interferencia de la información semántica que podría alterar la interpretación de las respuestas en las tareas que las involucran. En particular, dado que estos experimentos fueron realizados en distintas poblaciones lingüísticas, permitieron constatar la universalidad de las respuestas concurrentes.

arbitrariedad, debería ser considerada un rasgo de “diseño” de los lenguajes naturales (Perniss *et al.*, 2010; Sidhu y Pexman, 2017b). La confianza en el volumen de evidencias que favorecen esta nueva visión acerca del lenguaje ha llevado a afirmar que estamos superando la “narrativa de marginalización” acerca de los íconos verbales (Dingemanse, 2018) e ingresando en la era del “retorno de la no arbitrariedad” (Dingemanse *et al.*, 2015).

En lo que respecta al canal vocal-auditivo, el primer “principio de diseño” según Hockett, este tiene una estrecha vinculación con la arbitrariedad. Esta relación se apoyaba en el supuesto generalizado de que las vocalizaciones tienen *per se* menos potencial icónico que otro tipo de signos, como los gestos o las señas (Tomasello, 2008), por lo cual “el habla es por fuerza arbitraria” (Hockett, citado por Perlman, 2017, p. 386). Ahora bien, por una parte, parece evidente que la modalidad manual, propia de los gestos y de los lenguajes de señas, favorece la representación icónica de una amplia variedad de propiedades, tales como acciones, movimientos, formas y relaciones espaciales, dado que sus “articuladores” se manifiestan en un espacio tridimensional mediante la configuración, el movimiento y localización de las manos, y otros movimientos corporales. Por contraste, el carácter unidimensional y secuencial, que es propio de la modalidad vocal, pareciera tener más limitaciones para la iconicidad. En síntesis, se creía que mientras la iconicidad era un rasgo distintivo del contenido semántico y la función semiótica de los gestos y de los lenguajes de señas, estaba por fuerza prácticamente ausente en los lenguajes verbales. Sin embargo, nuevamente, las intuiciones consistentes de los hablantes sobre los significados de palabras o pseudopalabras, del mismo tipo que las constatadas por Sapir sobre la asociación entre ciertas vocales (/i/ y /a/) en una palabra

y el *tamaño* diferente de los objetos referidos (pequeño y grande, respectivamente), y por Köhler entre ciertas vocales y consonantes (en las pseudopalabras “maluma” vs. “takete”) y la *forma* de los objetos (redondeados vs. con líneas quebradas y ángulos, respectivamente), así como a nivel de la prosodia, la gramática, la sintaxis, etc. (Perniss *et al.*, 2010), desmienten la tesis de la pobre potencialidad icónica de la modalidad vocal¹⁷. Por otra parte, actualmente se comprende mejor la influencia diferenciada de las propiedades modales distintivas de ambos tipos de “articuladores” específicos, sonidos y movimientos corporales, y, por lo tanto, también cuáles rasgos tienden a ser más icónicos en uno u otro tipo de signos, en los lenguajes en los que predominan. Se han estudiado también sus combinaciones, ya que los gestos y las señas pueden ser sonoros, así como las vocalizaciones pueden tener componentes visuales. A su vez, algunos rasgos pueden ser igualmente bien referidos mediante signos icónicos en ambas modalidades, sonora y visual; por ejemplo, la estructura temporal de los eventos, la repetición, el tamaño (Dingemanse *et al.*, 2015). Como consecuencia de todo ello, no solo se ha desdibujado la tesis de un claro contraste en cuanto al potencial icónico de ambas modalidades, sino que, antes bien, los estudios recientes se orientan a explorar los potenciales icónicos distintivos derivados de la influencia específica de cada una de las modalidades, así como de

¹⁷ Uno de los argumentos más sólidos en contra de la hipótesis de los orígenes vocales del lenguaje frente a la hipótesis *gesture-first* (Tomasello, 2008) se basa en las aparentes limitaciones de la modalidad vocal para producir signos icónicos. Pero el simbolismo sonoro desmiente esta tesis, tanto que para algunos debería considerársela más bien un mito (Perlman, 2017). En consecuencia, resulta razonable pensar en la contribución no solo de los gestos sino también de las vocalizaciones icónicas para explicar el origen de las lenguas (Imai y Kita, 2014; Perlman, 2017; Perniss *et al.*, 2010). Este es otro de los efectos señalados al comienzo, que derivan de reconocer la ubicuidad del simbolismo sonoro en el lenguaje.

sus combinaciones, en los lenguajes vocales y manuales (Perniss, Özyürek y Morgan, 2015). Mientras que las señas pueden, por ejemplo, representar mejor las acciones, las relaciones espaciales y la forma visual, las vocalizaciones, por su parte, pueden enfatizar cualidades y expresar emociones a través de la prosodia (Dingemanse *et al.*, 2015; Perlman y Cain, 2014). En lo que refiere a su distribución cuantitativa, un primer estudio comparado reveló notables semejanzas en los porcentajes de iconicidad en similares tipos léxicos entre dos lenguas de señas (estadounidense y británica) y dos lenguas orales (inglés y español) (Perlman *et al.* 2018).

Este cambio de enfoque acerca del impacto de las modalidades sobre la iconicidad en ambas lenguas se apoya en otro previo acerca de las lenguas de señas, de índole más general. Antes del primer intento de caracterizar a la lengua de señas como un lenguaje, que se debe a Stokoe (1960), se interpretaba que las señas no eran más que pantomimas corporales o gestos puramente icónicos. Muy probablemente esta siga siendo la opinión corriente de quienes solo empleamos lenguas orales. Sin embargo, existe un generalizado acuerdo entre los especialistas en reconocer en estas lenguas una estructura lingüística equivalente a la de las lenguas habladas: estructura fonológica, morfológica y sintáctica, dialectización, prosodia, cambio histórico y aprendizaje por parte de los niños, así como funciones similares a la de las palabras en la modalidad oral, dadas sus propiedades sintácticas y semánticas (cfr. Goldin-Meadow y Brentari, 2017). Este cambio de visión ha llevado a aceptar más naturalmente que también en las lenguas de señas se verifica un cierto balance entre señas icónicas y señas arbitrarias (Perniss *et al.*, 2010). De este modo, se cuestionaron fuertemente dos de los “principios de diseño” propuestos por Hockett para caracterizar a todo

lenguaje humano, que resultaban aparentemente menos controversiales que otros, y que excluían a las lenguas de señas: la utilización del canal vocal-auditivo y la arbitrariedad¹⁸. Actualmente se piensa que es más apropiado caracterizar a las señas como signos en algunos casos y como gestos, propiamente dichos, en otros, manteniendo una distinción entre ambos, aun cuando sea borrosa en ocasiones (cfr. Perniss *et al.*, 2015), del mismo modo que el uso de la modalidad oral está también integrado por gestos no verbales, además de los que podrían llamarse “gestos verbales”, es decir, palabras cuya significación es tan icónica o deíctica como puede serlo la de los gestos correspondientes (Goldin Meadow y Brentari, 2017).

Por todo ello, tanto para el lenguaje verbal como para el lenguaje de señas resulta evidente la necesidad de contar con una caracterización “no clásica” del lenguaje humano, al menos si se toman en cuenta solo estas dos características: iconicidad e independencia de una particular modalidad. En síntesis, si además de la iconicidad en el lenguaje vocal, se advierte que todo sistema lingüístico, sea oral o de señas, incluye la utilización de una variedad de gestos, con funciones semióticas muchas veces complementarias a la de los signos lingüísticos, entonces es adecuado caracterizar al lenguaje como un *sistema semiótico heterogéneo* (Kendon, 2014) y *multimodal* (Vigliocco *et al.*, 2014), más que como un sistema homogéneo y monomodal o, incluso, considerando sus otros rasgos estructurales, como un sistema amodal. Volveremos sobre este punto más abajo.

¹⁸ Como observan Goldin-Meadow y Brentari (2017), en la última formulación de los principios de ‘diseño’ de 1978, Hockett ya no hace referencia a la utilización del canal vocal-auditivo, quizás por haberse convencido de que la modalidad no es determinante.

Para apreciar más fácilmente el papel de la iconidad en las lenguas orales, la examinaremos, en lo que sigue, solamente como una *propiedad general del léxico*. Sin embargo, como ya hemos señalado, no debe asumirse que solo son icónicas las palabras, y menos aún solo cierto tipo de palabras; por otra parte, tampoco debe asumirse que todos los componentes de una palabra son icónicos, ya que, en muchos casos, solo lo son algunos elementos que las componen; y, por último, no se debe suponer que todas las palabras icónicas lo son en el mismo grado, ni que su iconicidad está relacionada con el mismo tipo de información o contenido semántico. Estas precisiones están reflejadas, entre otras, en la sofisticada teoría semiótica de Peirce: la integración en un mismo signo o conjunto significativo de signos, y en un lenguaje en general, de los tres tipos semióticos, de acuerdo con las diferentes relaciones con los objetos de los que ellos son signos. Es así como todos los signos lingüísticos poseen rasgos icónicos, indíxicos y simbólicos, con grados variables de abstracción respecto a las propiedades involucradas en cada caso y en proporciones variables en las diferentes unidades lingüísticas y en cada lenguaje, con un “ideal” apuntado hacia el equilibrio entre las tres funciones semióticas¹⁹ (Nöth, 2015). La definición canónica de signo icónico se debe a Peirce: un signo cuya relación con su objeto “consiste en una mera semejanza” (C.P. 3.362). La semejanza nunca es identidad, y admite distintos tipos (*imágenes, diagramas y metáforas*) y grados de semejanza, por lo tanto, la iconicidad es, por definición peirceana, una propiedad no homogénea y

¹⁹ “[...] los signos más perfectos de todos son aquellos en los cuales los caracteres simbólicos, indicativos e icónicos están mezclados en proporciones tan equiparables como sea posible” (CP 4.4448).

gradual. Las ideas de Peirce incluyen muchas otras sutilezas, las que se han recogido en distintos paradigmas para el estudio de la iconicidad (cfr. Nöth, 2015).

Se han denominado “ideófonos” (también “expresivos” o “miméticos”) a las palabras sonoro-simbólicas que tienen relaciones de *semejanza directa* con sus significados (Dingemanse, 2012). Esta clase léxica tiene propiedades sintácticas, fonológicas y semánticas distintivas: evocan imágenes sensoriomotoras de manera vívida (Winter *et al.*, 2017). No obstante, los significados de estas palabras también varían en su iconicidad en una jerarquía de mayor a menor iconicidad cuando evocan sonidos, movimientos, patrones visuales, otras propiedades sensoriales y, finalmente, estados afectivos o psicológicos (Dingemanse, 2012). En muchas familias de lenguas, los ideófonos son numerosos (vgr. el japonés, otras lenguas asiáticas, y también lenguas africanas y americanas). Entre ellas, las palabras más directa y claramente icónicas son las *onomatopeyas* y las *interjecciones*, consideradas durante mucho tiempo como residuos exóticos o una “fracción despreciable”, y, por lo tanto, de “importancia secundaria”, en las lenguas naturales (Newmeyer, 1992). Las onomatopeyas imitan sonidos del entorno, por ejemplo, de acontecimientos, como “¡boom!” imita el sonido de una explosión, o de entidades, por ejemplo, las vocalizaciones de los animales, como “guau” imita el sonido del ladrido de un perro, o “miau”, el maullido de un gato, etc.²⁰. Las interjecciones, por su parte, derivan de vocalizaciones emocionales, por ejemplo, “¡uy!” (susto, asombro), “¡ohhh!” (sorpresa), “¡ay!” (dolor), etc. Nótese que, en este caso, el

²⁰ Por su parte, las diferencias existentes en las onomatopeyas de diferentes lenguas se explican por los diferentes inventarios fonológicos de las lenguas, así como, en ocasiones, por estar referidas a distintos rasgos perceptuo-motores de los referentes (Dingemanse *et al.*, 2015).

mapeo sonido-significado es monomodal: sonido-sonido, por lo que algunos autores han propuesto que, aunque sean más claramente icónicas que otro tipo de palabras, no deberían considerarse palabras sonoro-simbólicas, ya que estas solo deberían incluir correspondencias transmodales entre la forma y el significado. Nótese también que cuando el mapeo es modomodal, es más transparente y, por lo tanto, más rápidamente comprendido. Otro patrón común en los ideófonos son las formas sonoras repetidas reflejadas en la estructura de la palabra, para representar eventos repetidos o una estructura iterada de eventos (por ejemplo, “zigzag” o “tictac”)²¹, así como también el contraste entre vocales (vgr. [i]:[a]) para representar contrastes en intensidad o tamaño. Los patrones prosódicos también pueden mapear patrones de la experiencia (por ejemplo, cuando alargamos la pronunciación de una vocal para indicar tamaño o duración temporal, vgr. “graaaaande” o “leeeeento”). Distintos estudios han constatado que los significados de estas palabras son detectados por sujetos que no tienen conocimiento de las lenguas a las que pertenecen con un porcentaje superior al azar (Dingemanse *et al.*, 2016).

Una de las comprobaciones más interesantes, realizadas con hablantes competentes y con sujetos que desconocen las lenguas sobre cuyas palabras tienen que conjutar significados, así como frente a pseudopalabras, es que la iconicidad no es una propiedad binaria, sino que

²¹ La reduplicación se manifiesta a nivel de fonemas o morfemas, palabras, estructuras sintácticas y actos de habla. A nivel sintáctico, por ejemplo, son casos de reduplicación “iconicidad secuencial”, como en la expresión “*Veni, vidi, vici*”, o la “iconicidad de contigüidad”, en la que las palabras pertenecientes a un mismo dominio conceptual van juntas (Perniss y Vigliocco, 2014). La primera ejemplifica el tipo “diagramático” de iconicidad “horizontal”, donde la relación secuencial entre los elementos en la oración es icónica respecto de la relación secuencial entre los eventos referidos.

se manifiesta en grados variables (Vigliocco *et al.*, 2014). Los estudios sobre las lenguas de señas han proporcionado una primera taxonomía de señas icónicas que puede aplicarse también a las palabras habladas: las que tienen una forma de *iconicidad directa y transparente*; las que, en cambio, manifiestan *iconicidad indirecta o relativa*; y, finalmente, las que presentan una *iconicidad translúcida y oscura*, menos clara para los hablantes (Perniss y Vigliocco, 2014). Se sabe, además, que al convencionalizarse, las palabras sufren cambios a través del tiempo que oscurecen gradualmente su motivación icónica, por lo que palabras originalmente más transparentes pueden tornarse luego apenas translúcidas. Del mismo modo, esta iconicidad se hace menos perceptible en las palabras en las que se combinan rasgos icónicos y arbitrarios (Dingemanse *et al.*, 2016; Sidhu y Pexman, 2017b).

Adicionalmente, se observan grados variables de iconicidad en relación con las modalidades, con los dominios semánticos involucrados y con las clases léxicas. Como se ha dicho arriba, cuando los sonidos se correlacionan con sonidos, la iconicidad es mayor que cuando se correlacionan con propiedades visuales²², y así con las siguientes modalidades sensoriales, de las cuales las menos icónicas son aquellas relacionadas con el olfato, y en menor medida aún, o más indirectamente, las relacionadas con rasgos afectivos o emocionales (Dingemanse, 2012; Winter *et al.* 2017). Del mismo modo, se ha constatado que son más icónicas las palabras que refieren a propiedades sensoriales o que tienen contenido experiencial que aquellas que refieren a contenidos más abstractos, en una gradación que

²² Esta jerarquía se modifica en las lenguas de señas, donde las palabras más icónicas están asociadas a significados referidos a propiedades visuales, como el movimiento, la forma y las relaciones espaciales (Vinson, Cormier, Denmark, Schembri y Vigliocco, 2008).

cuantifica a las palabras entre las más y las menos icónicas (Lupyan *et al.*, 2018; Perry, Perlman y Lupyan, 2015; Sidhu y Pexman, 2017a; Winter *et al.*, 2017)²³. Por otra parte, también se ha estudiado el grado de iconicidad en distintos tipos léxicos. Un estudio comparado de unas 600 palabras en inglés y español mostró que los adjetivos son más icónicos que los nombres y las palabras funcionales, y que los verbos son más icónicos en inglés que en español, lo que podría obedecer a que en inglés incorporan información relativa al movimiento, que en el español queda reservada a otros elementos en la estructura oracional (Perry, Perlman y Lupyan, 2015). Otro estudio, realizado con unas 3.000 palabras en inglés y con hablantes nativos, corroboró esos datos, identificando grados de mayor a menor iconicidad en adjetivos y verbos, y luego en nombres y otros tipos de palabras (Winter *et al.*, 2017). Se ha comprobado, además, que otro factor que modula el grado de iconicidad en el léxico es el tipo de distribución que las palabras tienen en el “espacio semántico” de una lengua. En efecto, las palabras tienen una densidad semántica variable (*semantic neighbourhood density*), es decir, tienen entre sí contenidos semánticos similares en diferentes grados. Dado que el riesgo de confusión entre palabras con menos significados similares se minimiza, estas palabras pueden “aprovechar más las ventajas de la iconicidad” que las que tienen mayor densidad semántica (Sidhu y Pexman, 2017b). De ese modo se puede explicar que las palabras con menor densidad semántica tiendan a ser más icónicas.

²³ Similares resultados se verifican para los así llamados “fonestemas”, que son fonemas o conjuntos de fonemas al inicio o al final de las palabras, que portan información semántica, tales como /gl/, en palabras que refieren a rasgos visuales brillantes o relacionados con la luz y la visión, o /sn/ en palabras con significados nasales, en inglés, entre tantos otros. Los fonestemas serían la unidad mínima significativa de carácter icónico, entre los fonemas y los morfemas (cfr. Perniss *et al.*, 2010).

Así como no todas las palabras tienen el mismo grado de iconicidad, lo mismo puede decirse acerca de los lenguajes: hay lenguajes menos icónicos, como los indoeuropeos (el inglés, por ejemplo), y otros más icónicos, como las lenguas africanas subsaharianas, las de los países orientales (como el japonés o el coreano), y las lenguas australianas y americanas aborígenes (como el quechua). Aunque estudios comparados recientes de distintas familias de lenguas comprobaron la presencia ubicua de estos rasgos en todas las lenguas, incluidas las indoeuropeas, estos tienen diversa prevalencia y distribución en tipos diferentes de palabras. Estas variaciones podrían explicarse por diversos factores, entre ellos, los rasgos propios de cada lengua, especialmente su diferente repertorio fonológico, así como la intervención de otros factores que influyen en el cambio lingüístico a través del tiempo (Ahler y Zlatev, 2010; Blasi *et al.*, 2016; Dingemanse, 2012; Nielsen y Dingemanse, 2018; Perry, Perlman y Lupyan, 2015; Vigliocco *et al.*, 2014).

En cuanto a los mecanismos que podrían explicar cómo se originan las asociaciones sonoro-simbólicas, mencionaré brevemente los que se han conjeturado, cada uno con diversa base probatoria (Sidhu y Pexman, 2017b). Por una parte, la *co-ocurrencia estadística (?) de los estímulos* en el entorno o la relación entre estos y ciertas experiencias subjetivas son lo que explicaría que esas probabilidades luego sean internalizadas. Este tipo de mecanismo crearía asociaciones dependientes de la experiencia, que podrían ser compatibles con ciertas predisposiciones para adquirirlas. Este mecanismo también fue propuesto para las CT en general. Otro mecanismo sería el de las *propiedades compartidas* entre fonemas y demás rasgos fonológicos y algunos rasgos asociados: perceptivos (el tamaño, la forma, la duración), conceptuales, afectivos o lingüísticos,

en un espectro que varía entre vínculos más imitativos, donde las propiedades compartidas son más perceptivas, y vínculos más connotativos, donde son más conceptuales o lingüísticas. Este sería una variante del primer mecanismo y permitiría dar cuenta tanto de las congruencias sinestésicas como de las congruencias semánticas. Un tercer mecanismo sería el de las *propiedades estructurales del cerebro*. Según esta hipótesis, los “factores neurales” serían responsables de crear estas asociaciones. Esto daría por resultado asociaciones universales, como las identificadas por Sapir y Köhler. Otro mecanismo similar, y que podría considerarse una variante del anterior, serían las *asociaciones comunes a la especie* o a especies emparentadas, fruto de la evolución, que podría interpretarse en términos de predisposición evolutiva para adquirir ciertas asociaciones. También este fue propuesto para las CT. Los dos mecanismos anteriores explicarían las asociaciones sonoro-simbólicas en función de rasgos de naturaleza biológica. Finalmente, los *patrones lingüísticos* podrían ser responsables de crear estas asociaciones, o, más débilmente, de modular asociaciones preexistentes, por ejemplo, la recurrencia de ciertos fonemas en palabras con significados similares. También, se ha propuesto explicar las CT de manera similar. Por último, dada la diversidad de rasgos que manifiesta el simbolismo sonoro, la intervención de estos mecanismos podría no ser excluyente, es decir, podrían interactuar entre sí, o bien, en algunos casos, cada uno de ellos podría explicar diferentes variedades (Sidhu y Pexman, 2017b). Seguramente, estas distinciones, para las CT en general, y para el simbolismo sonoro en particular, se irán afinando y homogeneizando en el futuro. Pero puede conjeturarse que la variedad de tipos y grados de simbolismo sonoro en las

lenguas, en particular su carácter más o menos universal o cultural, es compatible con la intervención combinada de estos diferentes mecanismos (Dingemanse *et al.*, 2015).

4. Simbolismo sonoro y aprendizaje del lenguaje

Un tópico que ha contribuido mucho a evidenciar la peculiar importancia de la iconicidad en el lenguaje, reflejado en un creciente volumen de investigaciones empíricas (Nielsen y Dingemanse, 2018), es su relativa mayor presencia o frecuencia en el léxico que primero aprenden los niños, relativamente al léxico de los adultos (Perry *et al.*, 2015)²⁴, así como también en el léxico que los adultos emplean cuando se dirigen a los niños, el así llamado *maternés* (Imai y Kita, 2014; Perry *et al.*, 2017). Las palabras icónicas en el léxico permiten predecir la edad de su adquisición, es decir, son aprendidas antes, lo que sugiere que serían más fáciles de adquirir que las arbitrarias (Perry, Perlman y Lupyán, 2015). Además del léxico, los rasgos prosódicos distintivos del habla dirigida a los niños manifiestan mayor iconicidad que la que está presente en el habla dirigida a los adultos (Nygaard, Herold y Namy, 2009). Si se advierte que aprender un lenguaje es, al menos, aprender a mapear formas lingüísticas específicas con significados, parece bastante intuitivo suponer que las palabras más icónicas pueden ser asociadas con sus significados más fácil y rápidamente que las arbitrarias, y que ello explicaría el hecho de que el ingreso al lenguaje se vea

²⁴ La mayor parte de las palabras aprendidas por los niños son no arbitrarias hasta que, a los 6 años aproximadamente, comienza a incrementarse notablemente el volumen de palabras arbitrarias en el léxico. Según algunos estudios, solo superan en número a las icónicas aproximadamente a la edad de 13 años o más (Monaghan *et al.*, 2014).

facilitado por medio del aprendizaje y utilización de este tipo de palabras, basadas a su vez en capacidades previas, no aprendidas, para detectar las correspondencias transmodales en las que se basan (Imai y Kita, 2014; Monaghan *et al.*, 2014). Un mecanismo que permitiría reducir la ambigüedad referencial, acoplando de manera adecuada los sonidos con la referencia apropiada, sería justamente ligar los rasgos perceptuales de los primeros con el acceso sensorio-motor a los segundos. De ese modo, cuando los signos son icónicos es posible tender un puente entre la experiencia del mundo y la capacidad para comunicar esa experiencia (Perniss y Vigliocco, 2014). El mismo fenómeno se ha comprobado en las lenguas de señas (Vinson *et al.*, 2008). Por su parte, también se han encontrado diferencias en el aprendizaje según los tipos léxicos: los verbos pueden ser mejor aprendidos mediante asociaciones sonoro-simbólicas que los nombres. Se presume que la causa es que los verbos compiten menos unos con otros que los nombres, en tanto hacen distinciones menos finas entre ellos. La menor cantidad de verbos que de nombres en las lenguas corroboraría esta hipótesis (Imai y Kita, 2014). Del mismo modo se aprenderían los nombres de propiedades (adjetivos y adverbios), que son en proporción más sonoro-simbólicos que los nombres de entidades (sustantivos). Otra hipótesis, avalada por tests que comparan léxicos sonoro-simbólicos y arbitrarios, es que el simbolismo sonoro contribuiría más precisamente al aprendizaje de categorías sobre la base de las cuales pueden aprenderse luego significados más específicos (Brand, Monaghan y Walker, 2018). Una hipótesis más ambiciosa ha propuesto generalizar esta función del simbolismo sonoro al aprendizaje del lenguaje en general, considerando que tiene un rol de *bootstrapping* (Imai y Kita, 2014): los niños aprovecharían las capacidades innatas para detectar

algunas correspondencias transmodales, lo que les permitiría comprender gradualmente el papel referencial de los sonidos del habla, estableciendo así sus primeras representaciones léxicas sonoro-simbólicas. Sobre estas bases serían posibles ulteriores aprendizajes de otro tipo de palabras y construcciones en el lenguaje, en las que podrían servir de apoyo otras correspondencias aprendidas (Imai y Kita, 2014²⁵). Podría pensarse, por el contrario, que, en vez de servir como punto de apoyo o andamiaje sobre la base del cual es posible el aprendizaje ulterior de otro tipo de palabras, esta función estuviera limitada al repertorio específico de palabras icónicas, en cuyo caso el papel causal de la iconicidad en el aprendizaje general del lenguaje sería más acotado. La respuesta a esta cuestión sigue abierta, y depende aún de estudios y evidencias que respalden una u otra interpretación (Nielsen y Dingemanse, 2018). Por su parte, podría suponerse que las palabras más icónicas solo facilitarían una “comunicación más vívida de contenidos sensoriales” (Lupyan y Winter, 2018), antes que el aprendizaje del léxico respectivo. En tal caso, tendría igualmente una vinculación positiva con el aprendizaje de palabras, aunque más indirecta, en tanto el recurso a vocalizaciones icónicas para la comunicación consolida los escenarios interactivos que favorecen el temprano aprendizaje del lenguaje (Scotto, 2017a). En cualquier caso, la permanencia del vocabulario icónico en el lenguaje adulto,

²⁵ Sobre esta hipótesis, los resultados de algunos tests sugieren que, con el desarrollo, se produce un gradual estrechamiento de la sensibilidad sonoro-simbólica a los sonidos de habla de la propia lengua, aproximadamente a partir del año de vida. Ello explicaría que las primeras palabras sonoro-simbólicas fueran más universales, al estar basadas en los rasgos articulatorios del habla y los acústicos de los referentes, mientras que las adquiridas con posterioridad fueran específicas a las lenguas particulares, al estar limitadas al inventario fonológico de cada lengua, entre otros factores.

así como el uso de otros recursos icónicos en el lenguaje y la comunicación lingüística adulta, sugieren que su rol no está limitado al aprendizaje²⁶.

A la luz de estos resultados, la conocida solución chomskiana en favor de un conocimiento innato de estructuras universales para explicar el aprendizaje del lenguaje frente a la “pobreza del estímulo” debería ser reemplazada por una perspectiva diferente. Dado que las lenguas deben adaptarse a las demandas de aprendizaje de sus usuarios (Lupyan y Winter, 2018), los “aspectos aparentemente arbitrarios de la estructura pueden resultar del aprendizaje general y de sesgos en el procesamiento que derivan de la estructura de los procesos mentales, de factores perceptuo-motores, de limitaciones cognitivas y de constreñimientos pragmáticos” (Christiansen y Chater, 2008, p. 489), más que de una gramática universal subyacente que reserva la arbitrariedad al nivel de las herramientas comunicativas. Dicho de manera simplificada: dado que los lenguajes deben ser relativamente aprendibles según las capacidades cognitivas de las criaturas que serán sus usuarios, parecería razonable reemplazar la pregunta chomskiana “¿Por qué el cerebro está tan bien adaptado para aprender el lenguaje?” por la siguiente: “¿Por qué el lenguaje está tan bien adaptado para ser aprendido por el cerebro?” (Christiansen y Chater, 2008, p. 490). Según esta perspectiva, es útil dejar de ver *al lenguaje* como una estructura abstracta universal que hace aprendibles todas las lenguas, para ver a *los lenguajes*, en el sentido de “*las*

²⁶ Su permanencia e importancia relativa en todas las lenguas podría explicarse, entre otras hipótesis complementarias, por su capacidad para sugerir vocalizaciones nuevas y significativas (Perlman, Dale y Lupyan, 2015), es decir, el simbolismo sonoro podría constituirse en un mecanismo que facilita la formación de nuevas palabras (Dingemanse, 2014). De ese modo, podría contribuir a entender el cambio lingüístico y, por esa razón, también los orígenes del lenguaje (Imai y Kita, 2014).

lenguas", más bien "como 'organismos' que han evolucionado en una relación simbiótica con los seres humanos" (Christiansen y Chater, 2008, p. 490). Dicho de otra forma, se debería reemplazar la idea del lenguaje como un sistema estático ya "diseñado" (una noción cuanto menos ambigua) por la idea de las lenguas como sistemas dinámicos y complejos que han evolucionado.

Finalmente, ¿qué se sigue de las anteriores consideraciones acerca de la arbitrariedad en relación con el aprendizaje de palabras? En primer lugar, parece intuitivo que las palabras cuyo contenido semántico es más abstracto tendrían que estar disponibles para los niños una vez que ya disponen de un primer léxico más ligado a referentes y a contextos concretos, y a medida que se incorporan más plenamente a la cultura de su comunidad. Dicho de otra forma, los niños necesitan primero aprender a referirse a las *manzanas*, y solo mucho más tarde a la *justicia*. Parece importante, en consecuencia, explorar la distinción entre *palabras concretas* y *abstractas* en términos del conocimiento semántico involucrado, para determinar, a su vez, en qué medida iconicidad y arbitrariedad se distribuyen en el léxico y en el aprendizaje. Las palabras concretas han sido definidas como aquellas "que refieren a cosas o acciones en la realidad que puedes experimentar directamente a través de uno de los cinco sentidos" (Brysbaert, Warrinen y Kuperman, 2014), o a manifestaciones físicas directamente perceptibles. La mayor parte de las palabras icónicas deberían estar entre ellas. Por contraste, las palabras abstractas se definen como aquellas que refieren a "significados que no pueden ser experimentados directamente, pero que conocemos porque [...] pueden ser definidos por medio de otras palabras" (Brysbaert *et al.*, 2014, p. 906). La arbitrariedad parecería más apropiada para este tipo de palabras. Un test realizado sobre una

muestra de 40.000 palabras de distintos tipos léxicos del inglés, en el que los participantes debían distinguir palabras concretas de abstractas, reveló una mayor proporción de palabras con contenidos abstractos, aunque en grados variables, sobre palabras concretas. Este test permitió valorar ambos tipos de palabras en un espectro que va de 1, para palabras más abstractas, a 5, para palabras más concretas. Incluso los nombres comunes, que son las palabras que pertenecen a “las clases léxicas más concretas de todas”, fueron juzgadas como más abstractas, en promedio, que otras palabras canónicamente abstractas (vgr. *fun* con 1,97 vs *freedom* con 2,34). Considerando otro tipo de palabras, como los verbos, los adverbios y los adjetivos, de uso corriente, se advirtieron también importantes grados de abstracción, que superan los que podrían presumirse *a priori*. Aunque no se conocen estudios similares sobre otras lenguas, estos resultados ponen en evidencia no solo la ubicuidad de las palabras abstractas en el inglés, sino que incluso aquellas no definidas como tales son también evaluadas por los usuarios como abstractas (Lupyan y Winter, 2018). Ello explicaría que la mayor parte de las palabras, al menos en inglés, tengan vínculos arbitrarios y no icónicos con sus significados. Sin embargo, algunas precisiones podrían ser útiles para interpretar adecuadamente esta conclusión. Aun siendo ilustrativa de un contraste real claramente ejemplificado por palabras muy concretas frente a palabras muy abstractas, con gradientes de abstracción intermedios, esta distinción no sería coextensiva con la distinción entre arbitrariedad e iconicidad, en cuanto las semejanzas de los signos con sus significados que generan iconicidad también pueden ser muy indirectas y esquemáticas, y, por lo tanto, no deberían limitarse al caso de las palabras que solo refieren a rasgos sensoriales de los objetos o eventos. Por otra parte, se debe tomar en

cuenta que una misma palabra puede ser en parte icónica (o sistemática) y en parte arbitraria, por lo que en muchos casos es difícil clasificar el repertorio léxico para evaluar iconicidad frente a arbitrariedad solo sobre la base de esta distinción. La conclusión de este estudio es que el lenguaje es menos icónico que lo que podría ser, y que este es un fenómeno que requiere explicación. Es bastante intuitiva, y es, además, dado el conocimiento de muchas lenguas, correcta. Sin embargo, su hipótesis explicativa, a saber, que la iconicidad conspira contra la abstracción, requeriría un examen de los conceptos y palabras abstractas más refinado del que se ha realizado hasta ahora (cfr. Barsalou, 2016). Otros resultados muestran que si la iconicidad fuera un rasgo dominante en el léxico, los lenguajes verbales perderían gran parte de su poder expresivo (Dingemanse *et al.*, 2015). En ese sentido, si las palabras similares por su forma tuvieran significados similares, podría generarse confusión (Monaghan *et al.*, 2014; Lupyan y Winter, 2018). La necesidad de evitar este problema impone un límite sobre la iconicidad diferente al impuesto por la abstracción. El problema se advierte más claramente cuando el tamaño del vocabulario se incrementa (Dingemanse *et al.*, 2015), y se manifiesta, como vimos, en el hecho de que palabras con menos “vecinos semánticos” tienen mayor iconicidad que aquellas con mayor cantidad de palabras semánticamente emparentadas (Sidhu y Pexman, 2017b). La arbitrariedad puede evitar que las semejanzas semánticas generen confusión al ser vehiculizadas por signos también semejantes, es decir, permite que palabras con significados similares sean fonológicamente diferentes para minimizar la interferencia y la confusión (Lupyan y Winter, 2018). Podría pensarse, entonces, que, si la iconicidad ofrece las ventajas conjeturadas en la facilitación del aprendizaje temprano de las primeras palabras, así como, probablemente,

en dar origen a los lenguajes verbales, esa función es gradual y parcialmente reemplazada por la arbitrariedad, en proporciones variables según las lenguas. Este fenómeno, sin embargo, es también compatible con el hecho de que los rasgos icónicos originales, además de combinarse con rasgos arbitrarios, no hayan desaparecido del todo, sino que solo se hayan tornado menos transparentes para los hablantes (Imai y Kita, 2014).

Por último, dadas las funciones diferentes que arbitrariedad e iconicidad desempeñan en el lenguaje, otro tópico de gran interés es cuál sería el modelo adecuado para explicar el *conocimiento semántico* de ambos tipos de palabras. Esta cuestión está también estrechamente relacionada con la caracterización que se adopte de los rasgos definitorios de los lenguajes humanos. Aunque ha sido planteada en el nivel más general de las representaciones conceptuales, en términos de las características del formato de los conceptos, podemos abordarla teniendo en miras el conocimiento del significado lingüístico. Por una parte, el enfoque de la *cognición corporizada*, o, más ampliamente, de la *cognición grounded*²⁷, afirma que todos los procesos cognitivos están enraizados en la percepción y la acción (Meteyard, Rodríguez Cuadrado, Bahrami y Vigliocco, 2012). Para estas teorías, hay un isomorfismo entre el formato de una representación conceptual y su contenido, es decir, el formato representacional no es arbitrario, y, por lo tanto, los conceptos están representados por medio de información sensorio-motora, es decir, son modales (Mahon y Hickok, 2016). Un enfoque como este podría explicar mejor el conocimiento de conceptos concretos e icónicos sobre la base de experiencias sensoriales, motoras

²⁷ Este enfoque incluye, además del cuerpo y las modalidades, al entorno físico y social, por lo tanto, es un enfoque más amplio que el corporizado (Barsalou, 2016).

y afectivas, aunque debería intentar explicar los restantes sobre la base de aquellos o bien mediante la intervención combinada de otros mecanismos. Todo ello supone que el sistema sensorio-motor y afectivo esté involucrado en el procesamiento del significado de los signos lingüísticos (Perniss *et al.*, 2010; Perniss y Vigliocco, 2014). Por otra parte, el *enfoque amodal*²⁸ o *simbólico* (o simplemente “clásico”) concibe al lenguaje mismo como un sistema completamente simbólico y abstracto, cuyos símbolos son procesados con independencia de su contenido, y, por lo tanto, poseen un formato liberado de cualquier contenido modal (Meteyard *et al.*, 2012). Estas teorías defienden que los conceptos están representados como símbolos amodales y que estos no están constreñidos por el contenido de los conceptos (Mahon y Hickok, 2016). Dado que las palabras representan conceptos amodales, sus significados también lo son. El sistema de procesamiento semántico/gramatical, de acuerdo con este enfoque, está en un nivel distinto al del procesamiento fonológico, constituyendo ambos dos sistemas paralelos e independientes. En principio, este enfoque podría explicar mejor los conceptos abstractos, aquellos que mantienen una relación arbitraria con sus contenidos, sin embargo, pretende explicar todo tipo de conceptos o significados. Es también cierto que la adhesión a este enfoque depende, en gran medida, de la generalizada aceptación de la tesis de la arbitrariedad entre signos y significados (Mahon y Hickok, 2016). Por la misma razón, tiene dificultades para explicar adecuadamente las correspondencias sonoro-simbólicas presentes

²⁸ En este contexto, dependiendo de los autores, “amodal” significa ‘simbólico’ o ‘abstracto’ en cuanto opuesto a “íconico”; independiente del *input* perceptual o de las representaciones perceptuales; no isomórfico, en el sentido de que el vehículo conceptual no se corresponde con (o no se asemeja a) los rasgos del contenido representado, o que es arbitrario en relación con el contenido que representa.

en todos los lenguajes verbales. Ahora bien, si el lenguaje fuera concebido como un sistema semiótico heterogéneo, en el sentido de que no todas las palabras se relacionan con sus significados de la misma manera, podría pensarse en la pertinencia de combinar ambos enfoques, o, dicho de otra forma, de ver un continuo de posibilidades teóricas entre un enfoque corporizado y multimodal fuerte y otro igualmente fuerte, completamente abstracto y amostral, identificando, entre ambas alternativas teóricas, versiones menos fuertes de uno y otro signo, más compatibles con la heterogeneidad de las representaciones lingüísticas (y conceptuales) que se intentan explicar (Meteyard *et al.*, 2012). Esta es una conclusión similar a la sugerida por Barsalou, cuando afirma que “un desafío actual es explicar cómo *grounding* y abstracción emergen juntos” (Barsalou, 2016, p. 1132). Se trata, por cierto, de una discusión meta-teórica muy compleja y aún abierta, que excede ampliamente los límites de este trabajo, pero que, como es evidente, está directamente vinculada con una respuesta a la pregunta: “¿Cómo se relacionan los signos con sus significados en los lenguajes verbales humanos?”, y a la más específica “¿En qué consiste conocer el significado de distinto tipo de palabras?”. Ambas preguntas, a su vez, son relevantes por su impacto sobre una pregunta más fundamental: “¿Qué clase de entidad es el lenguaje humano?”.

5. Algunas conclusiones y cuestiones abiertas

El examen del simbolismo sonoro permite extraer algunas conclusiones respecto a la naturaleza de los lenguajes verbales que vale la pena precisar, para no confundir ni su contenido ni su alcance. La presencia de rasgos icónicos, en diversas variedades y grados, pero en mayor cantidad y

variedad que lo que se había reconocido tradicionalmente, en todos los lenguajes estudiados, permite afirmar que la iconicidad es una *propiedad general del lenguaje*. Sin embargo, esta conclusión no equivale a afirmar que hay en el lenguaje más iconicidad que arbitrariedad. La conclusión correcta sería, más bien, que hay más iconicidad que la que se había considerado tradicionalmente. Esta conclusión está directamente ligada a otra que, en parte, la hace comprensible: hay más rasgos icónicos que los que resultan evidentes o transparentes para los propios hablantes (Monaghan *et al.*, 2014). Dicho de otra forma, se trata de rasgos que, más allá de la sensibilidad semántica que la mayor parte de la población evidencia para detectarlos, no son identificados como tales sin esfuerzo reflexivo o conocimiento teórico, por algunas de las razones que hemos referido en este trabajo. Esta conclusión se complementa, entonces, con otra: iconicidad y arbitrariedad, en cuanto propiedades generales del lenguaje, desempeñan distintas funciones, y en tal sentido podrían ser adaptaciones específicas a distintos constreñimientos del lenguaje: mientras que la iconicidad procura “vincular la forma lingüística con la experiencia humana”, la arbitrariedad procura satisfacer “la necesidad de que la señal lingüística sea eficiente y discriminable” (Perniss y Vigliocco, 2014). Por lo tanto, ambas coexisten y se complementan mediante una lógica de “división del trabajo” en todos los niveles señalados (Christiansen y Monaghan, 2015), o, si se prefiere, ambos rasgos conviven en el lenguaje mediante una “síntesis dialéctica” (Monaghan *et al.*, 2014). En suma, la *ubicuidad* del simbolismo sonoro se manifiesta en su presencia en todos los niveles y dimensiones del lenguaje (aquí solo hemos analizado el nivel léxico) y en todas las lenguas estudiadas. En cuanto a su *importancia relativa*, se deriva tanto de su

carácter no marginal en cada lengua, como de las diversas funciones semióticas y cognitivas que satisface (en este trabajo nos hemos referido sobre todo al aprendizaje).

En términos más generales, la presencia ubicua y la importancia relativa de las CT bajo la forma del simbolismo sonoro, al poner en evidencia los vínculos constitutivos entre percepción y lenguaje, favorecen la propuesta de un *enfoque multimodal del lenguaje* (Vigliocco *et al.*, 2014). Por contraste, estos vínculos y los rasgos modales del conocimiento lingüístico son desplazados hacia la “periferia” en el paradigma ortodoxo. El cambio de enfoque que deriva de la importancia de los rasgos icónicos en el lenguaje verbal se apoya en otros argumentos empíricos complementarios, especialmente la presencia y el papel de la gestualidad en la comunicación lingüística. Siendo la iconicidad una propiedad general del sistema lingüístico, integrado por signos lingüísticos y también por gestos y otros comportamientos expresivos, este deberá ser visto como un *sistema semiótico heterogéneo*. Los signos de este sistema son vehiculizados a través de distintas modalidades y dimensiones, tanto visual (gestual, corporal y facial) como vocal (prosódica y articulatoria), y se vinculan de maneras congruentes con sus significados, mediante una amplia variedad de otras modalidades sensoriales o con estados afectivos u otros fenómenos experienciales. De acuerdo con estos elementos, podría suponerse que la perspectiva teórica más adecuada para el estudio del lenguaje, al menos para dar cuenta de estos fenómenos icónicos y multimodales, es el enfoque de la *cognición corporizada* o “*grounded*”. Sin embargo, como también hemos señalado, es probable que, dados otros rasgos distintivos o propiedades generales del lenguaje, que son aquellos en los que más se interesaron lingüistas y filósofos, tales como las palabras abstractas referidas mediante signos arbitrarios, este

enfoque deba combinarse con alguna variante de *conceptión amodal* del lenguaje. Esta es una cuestión abierta. Precisar mejor y fundamentar acabadamente cada una de estas tesis, y sobre todo definir un enfoque integralmente apropiado para abordar el estudio del lenguaje, continúa siendo una tarea para las investigaciones empíricas y teóricas por venir. En lo que a la filosofía del lenguaje se refiere, no se necesita revisar la bibliografía más relevante para afirmar que, después de Peirce, quizás con la excepción de Wittgenstein²⁹, el fenómeno de la iconicidad en el lenguaje ha pasado casi completamente inadvertido. Aliento la expectativa de que este trabajo, que sobre todo ha intentado integrar algunos de los resultados empíricos más recientes sobre el tema, sea capaz de motivar el interés por un cambio de perspectiva.

Referencias bibliográficas

- Ahlner, F. y Zlatev, J. (2010). "Cross-modal iconicity: A cognitive semiotic approach to sound-symbolism". *Sign System Studies*, 38 (1/4), pp. 298-348.
- Barsalou, L. W. (2016). "On Staying Grounded and Avoided Quixotic Dead Ends". *Psychonomic Bulletin Review*, 23(4), pp. 1122-1142.
- Brysbaert, M., Warrinen, A. E. y Kuperman, V. (2014). "Concreteness ratings for 40 thousands generally known English lemmas". *Behav. Res. Methods*, 46, pp. 904-911.

²⁹ En "Wittgenstein: iconicidad en el lenguaje y 'experiencia del significado'" (2017b), desarrollo una hipótesis de lectura acerca del reconocimiento de distintos fenómenos icónicos en la obra tardía de Wittgenstein, donde también me refiero a varios tópicos que en este trabajo abordo *in extenso*.

- Blasi, D. E., Wichmann, S., Hammarström, H., Stadler, P. F. y Christiansen, M. H. (2016). "Sound-meaning association biases evidenced across thousands of languages". *Proceedings of the National Academy of Sciences*, 113, 10818–10823. DOI:10.1073/pnas.1605782113.
- Brand, J., Monaghan, P. y Walker, P. (2018). Changing Signs: Testing How Sound-symbolism Supports Early Word Learning. *CogSci 2018. Cognitive Science Society*. <https://goo.gl/49YDwK>.
- Christiansen, M. H. y Chater, N. (2008). "Language as shaped by the brain". *Behavioral and Brain Sciences*, 31, pp. 489-509.
- Christiansen, M.H. y Monaghan, P. (2015). "Division of Labor in Vocabulary Structure. Insights from Corpus Analyses". *Topics in Cognitive Science*, 8, 610-624. DOI: 10.1111/tops.12164.
- Deroy, O. y Spence, C. (2016). "Cross-Modal Correspondences: Four Challenges". *Multisensory Research*, 29, pp. 29-48.
- De Saussure, F. (1916). *Course in General Linguistics*. La Salle, IL: Open Court, 1983.
- Dingemanse, M. (2012). "Advances in the Cross-Linguistic Study of Ideophones". *Lang. Linguistic Compass*, 6, 654–672. DOI:10.1002/lnc3.361.
- Dingemanse, M., Blasi, D. E., Lupyan, G., Christiansen, M. H. y Monaghan, P. (2015). "Arbitrariness, Iconicity, and Systematicity in Language". *Trends in Cognitive Science*, 19(10), pp. 603-615.
- Dingemanse, M., Schuerman, W., Reinish, E., Tufvesson, S. y Mitterer, H. (2016). "What sound symbolism can and cannot do: Testing the iconicity of ideophones from five languages". *Language*, 92 (2), e-117-e133.

- Dingemanse, M. (2018). "Redrawing the margins of language: Lessons from research on ideophones". *Glossa: A Journal of General Linguistics*, 3(1), 4, pp. 1-30.
- Emmorey, K. (2014). "Iconicity as structure mapping". *Philosophical Transactions of Royal Society B* 369. DOI: 10.1098/rstb.2013.0301.
- Forceville, C. (2006). "Non-verbal and multimodal metaphor in a cognitivist framework: agendas for research". En G. Kristiansen, M. Achard, R. Dirven y F. Ruiz de Mendoza Ibáñez (Eds.), *Cognitive Linguistics: Current Applications and Future Perspectives* (pp. 379-402). NY: Mouton de Gruyter.
- Goldin-Meadow, S. y Brentari, D. (2017). "Gesture, sign and langue: The coming of age of sign language and gesture studies". *Behavioral and Brain Sciences*, 40, E46. DOI: 10.1017/S0140525X15001247.
- Hinton, L., Nichols, J. y Ohala, J. J. (1994). "Introduction: Sound-symbolism processes". En L. Hinton, J. Nichols y J. J. Ohala (Eds.), *Sound Symbolism* (pp. 1-12). Cambridge, England; New York, NY: Cambridge University Press.
- Hockett, C. (1960). "The Origin of Speech". *Scientific American*, 203, pp. 88-96.
- Hockett, C. (1963). "The problem of universals in language". En J. Greenberg. (Ed.), *Universals of Language* (pp. 1-22). Cambridge, MA: M.I.T. Press.
- Imai, M. y Kita, S. (2014). "The sound symbolism bootstrapping hypothesis for language acquisition and language evolution". *Philosophical Transactions of the Royal Society B*, 369: 20130298.
- Jakobson, R. (1965). "Quest for the Essence of Language". En *Selected Writings II, Word and Language* (pp. 345-359). The Hague: Mouton, 1971.

- Jespersen, O. (1921). *Language: Its Nature, Development and Origin*. London: George Allen and Unwin.
- Kendon, A. (2014). "Semiotic diversity in utterance production and the concept of 'language'". *Philosophical Transactions of the Royal Society B* 369, no. 1651: 20130293. DOI: 10.1098/rstb.2013.0293.
- Köhler, W. (1929). *Gestalt Psychology*. New York: Liveright.
- Locke, J. [1690], *Ensayo sobre el entendimiento humano*, México: FCE, 1956.
- Lupyan, G. y Winter, B. (2018). "Language is more abstract than you think, or, why aren't languages more iconic?". *Philosophical Transactions of the Royal Society B*, 373: 20170137.
- Lupyan, G. y Jerry, M. (2017). "From words-as-mappings to words-as-cues: the role of language in semantic knowledge". *Language, Cognition and Neuroscience*. DOI: 10.1080/23273798.2017.1404114.
- Mahon, B. Z. y Hickok, G. (2016). "Arguments about the nature of concepts: Symbols, embodiment, and beyond". *Psychonomic Bulletin y Review*, 23(4), pp. 941-958.
- Martino, G. y Marks, L. E. (2001). Synesthesia: Strong and Weak. *Curr.Dir. Psychol. Sci.* 10, 61-65. DOI: 10.1111/1467-8721.00116.
- Macpherson, F. (2011). "Individuating the senses". En F. Macpherson (Ed.), *The Senses: Classical and Contemporary Philosophical Perspectives* (pp. 3-43). Oxford: Oxford University Press.
- Meteyard, L., Rodríguez Cuadrado, S., Bahrami, B. y Vigliocco, G. (2012). "The coming of age: A review of embodiment and the neuroscience of semantics". *Cortex* 48, pp. 788-804.

- Monaghan, P., Shillcock, R. C., Christiansen, M.H. y Kirby, S. (2014). "How arbitrary is language?" *Philosophical Transactions of the Royal Society B* 369: 20130299. <https://goo.gl/aWbPf3>.
- Newmeyer, F. J. (1992). "Iconicity and Generative Grammar". *Language*, 68(4), pp. 756-796.
- Nielsen, A. K. y Dingemanse, M. (2018). "Iconicity in word learning and beyond: A critical review". *PsyArXivPre-prints*. <https://goo.gl/o89wEF>.
- Nöth, W. (2015). "Three paradigms of iconicity research on language and literature". En K. Hiraga Masako, W. J. Herlofsky, K. Shinoara y K. Akita (Eds.), *Iconicity. East Meets West* (pp. 13-34). Amsterdam/Philadelphia: John Benjamin Publishing. DOI: 10.1075/ill.2.05not
- Nygaard, L.C., Herold, D.S. Namy, L.L. (2009). "The Semantics of Prosody: Acoustic and Perceptual Evidence of Prosodic Correlates to Word Meaning". *Cognitive Science*, 33, 127-146. DOI: 10.1111/j.1551-6709.2008.01007.x.
- O'Callaghan, C. (2012). "Perception and Multimodality". En E. Margolis, R. Samuels y S. Stich. (Eds.), *The Oxford Handbook of Philosophy of Cognitive Science* (pp. 92-117). Oxford: Oxford University Press.
- Palmer, S. E., Schloss, K. B., Xu, Z. y Prado-León, L. R. (2013). "Music-color associations are mediated by emotion". *Proc. Nat. Acad. Sci. U.S.A.* 110, 8836-8841. DOI: 10.1073/pnas.1212562110.
- Parise, C. V. y Spence, Ch. (2012). "Audio-visual crossmodal correspondences and sound-symbolism: a study using the implicit association test". *Experimental Brain Research*, 220, pp. 319-333.
- Parise, C.V. (2016). "Crossmodal correspondences: Standing issues and experimental guidelines". *Multisensory Research*, 29 (1-3), pp. 7-28.

- Peirce, C. S. [1965]. *El hombre, un signo*. Barcelona: Ed. Crítica, 1988.
- Perlman, M. y Cain, A. A. (2014). "Iconicity in vocalization, comparisons with gesture, and implications for theories on the evolution of language". *Gesture*, 14(3), pp. 320-350.
- Perlman, M., Dale, R. y Lupyan, G. (2015). "Iconicity can ground the creation of vocal symbols". *Royal Society Open Science*, 2, 150152. DOI: 10.1098/rsos.150152.
- Perlman, M. (2017). "Debunking two myths against vocal origins of language. Language is iconic and multimodal in the core". *Interaction Studies*, 18(3), pp. 376-401.
- Perlman, M., Little, H., Thompson, B y Thompson, R. L. (2018). "Iconicity in Signed and Spoken Vocabulary: A Comparison between American Sign Language, British Sign Language, English and Spanish". *Frontiers in Psychology*, 9. <https://goo.gl/2EUJYc>.
- Perniss, P., Thompson, R. L. y Vigliocco, G. (2010). "Iconicity as a general property of language: Evidence from spoken and signed languages". *Frontiers in Psychology*, 1, 227.103389/fpsyg.
- Perniss, P. y Vigliocco, G. (2014). "The bridge of iconicity: From a world of experience to the experience of language". *Philosophical Transactions of Royal Society Lond B Biol Sci*, 369(1651):20130300.
- Perry, L. K., Perlman, M. y Lupyan, G. (2015). "Iconicity in English and Spanish and its Relation to Lexical Category and Age of Acquisition". *PLoS One*, 10(9):e0137147.
- Perry, L. K., Perlman, M., Winter, B. y Massaro, D. W. (2017). Iconicity in the speech of children and adults. *Developmental Science*, e12572.
- Platón. *Crátilo*. Madrid: Ed. Gredos, 2003.

- Ramachandran, V. S. y Hubbard, E. M. (2001). "Synaesthesia - A Window into Perception, Thought and Language". *Journal of Consciousness Studies*, 8(12), pp. 3-34.
- Sapir, E. (1929). "A study in phonetic symbolism". *J. Exp. Psychol.*, 12, pp. 225-239.
- Scotto, C. (2017a). "Lo que el aprendizaje del lenguaje revela sobre el lenguaje (y sobre la cognición social)". En D. Pérez y D. Lawler (Eds.), *La segunda persona y las emociones* (pp. 87-149). Buenos Aires: Sadaf Ed.
- Scotto, C. (2017b). "Wittgenstein: iconicidad en el lenguaje y 'experiencia del significado'". *Disputatio. Philosophical Research Bulletin*, 6(7), pp. 423-457.
- Sidhu, D. M. y Pexman, P.M. (2017a). "Lonely sensational icons: semantic neighbourhood density, sensory experience and iconicity". *Language, Cognition and Neuroscience*, 33(1), pp. 25-31.
- Sidhu, D. M. y Pexman, P.M. (2017b). "Five mechanisms of sound symbolic association". *Psychonomic Bulletin y Review*, 1-25. <https://DOI.org/10.3758/s13423-017-1361-1>
- Simner, J. (2007). "Synaesthesia as a psycholinguistic phenomenon". *Trends in Cognitive Science*, 11(1), pp. 23-29.
- Spence, C. (2011). "Crossmodal correspondences: A tutorial review". *Attention, Perception, y Psychophysics*, 73, pp. 971-995.
- Stokoe, W. C. (1960). *Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf*. Buffalo, NY: Univ. Buffalo Sign Language Structure.
- Tomasello, M. (2008). *Los orígenes de la comunicación humana*. Buenos Aires: Katz Editores, 2013.

- Vigliocco, G., Perniss, P. y Vinson, D. (2014). "Language as a multimodal phenomenon: implications for language learning, processing and evolution". *Philosophical Transactions of the Royal Society B* 369: 20130292.
- Vinson, D. P., Cormier, K., Denmark, T., Schembri, A. y Vigliocco, G. (2008). "The British Sign Language (BSL) norms for age of acquisition, familiarity, and iconicity". *Behavior Research Methods*, 40, 1079–1087. DOI: 10.3758/BRM.40.4.1079
- Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A. y Johnson, S. P. (2010). "Preverbal infant's sensitivity to synesthetic cross-modality correspondences". *Psychological Science*, 21, pp. 21-25.
- Wan, X., Woods, A.T., Van Den Bosch, J. J. F., McKenzie, K.J., Velasco, C. y Spence, C. (2014). "Cross-cultural differences in crossmodal correspondences between basic tastes and visual features". *Frontiers in Psychology*, 5, 1365. <https://goo.gl/nyzxJ9>.
- Winter, B., Perlman, M., Perry, L. K. y Lupyan, G. (2017). "What words are more iconic? Iconicity in English sensory words". *Interaction Studies*, 18(3), pp. 430-451.

2

Señalar el tiempo

El gesto deíctico desde un enfoque multimodal del lenguaje

IRENE AUDISIO (CIFFyH – UNC)

1. Gestos y palabras

Durante décadas, en las líneas de investigación más difundidas acerca del lenguaje, los gestos fueron dejados de lado. Recordemos, por ejemplo, la hipótesis de los orígenes del lenguaje en las vocalizaciones de los primates no humanos, que parecieron asumir que la genealogía del lenguaje debía restringirse a una historia de las vocalizaciones, y el lenguaje mismo, a un sistema vocal (Seyfarth, 1987; Snowdon, Brown y Petersen, 1982). Por otra parte, desde la perspectiva que introdujeron los enfoques pragmáticos del lenguaje, si bien se integró el estudio de los gestos, se los desplazó a un papel contextual y claramente desligado de las funciones plenamente lingüísticas (McNeill, 1985; 2005).

A partir de desarrollos lingüísticos, psicológicos, etológicos y filosóficos más recientes (Goldin-Meadow, 2003; McNeill, 2005; Tomasello, 2008), el foco de las investigaciones se dirige a los gestos, atribuyéndoles incluso un rol constitutivo del lenguaje mismo, al equipararlos con los componentes verbales y ponerlos en una relación dia-

léctica de coexpresividad. El lenguaje es definido en esta dialéctica entre imagen y habla (*image-speech*). Se piensa en una sincronía de esos dos tipos semióticos diversos en la que se basa su coexpresividad. Los gestos proveen las imágenes, y, en la tensión dialéctica entre esos modos diversos, gestual y vocal, se dinamiza el habla e incluso el pensamiento. En este marco multimodal, la gestualidad es un componente que integra el lenguaje y no un mero acompañante modulador.

Siguiendo esta dirección, a partir de las caracterizaciones que ofrece McNeill (1992) sobre el rol activo de los gestos en el lenguaje y la dialéctica gesto-palabra, nos proponemos estudiar un caso clave que brinda evidencia a favor de este enfoque.

Entre los gestos estudiados hay un tipo que nos interesa analizar aquí por su relevancia en relación con los orígenes filogenéticos y ontogenéticos del lenguaje. El tipo fundamental de gesto que cumple la función de acto comunicativo completo es el gesto para orientar la atención o gesto deíctico, cuyo prototipo es el acto de señalar (Kita, 2003; Tomasello, 2008). El señalamiento tiene una importancia clave en el desarrollo de la comunicación y, por lo tanto, del lenguaje. Encarna el modo más simple de intencionalidad compartida, ya que presupone una actividad interpersonal. Mediante un movimiento se logra captar la mirada del otro y dirigirla hacia el objeto del entorno inmediato sobre el cual se desea llamar la atención. Este acto de referencia tiene como propósito conseguir que el receptor infiera la intención social del comunicador, es decir, lo que quiere que haga, sienta o piense (Tomasello, 2008, p. 54). Inferir la intención social requiere la capacidad cognitiva de averiguar por qué se ha realizado ese acto referencial y qué pretende el comunicador del receptor. Esta labor cognitiva es propia de la interpretación que requiere el gesto

deíctico pleno. Este carácter claramente comunicativo fue apareciendo y complejizándose evolutivamente, se presenta desde los primates hasta los humanos, incorporando cada vez grados más complejos de intencionalidad.

En el caso del señalamiento humano, las tres funciones que cumple el gesto deíctico sientan las bases de la comunicación. Se usa para compartir, informar y pedir: “[...] las tres actitudes implican de una manera u otra los móviles cooperativos de ayudar y compartir, que son los tipos principales de motivación subyacente en la intencionalidad compartida” (Tomasello, 2008, p. 97).

La importancia teórica del señalamiento ha sido usualmente abordada tanto desde los puntos de vista evolutivo filogenético como del desarrollo. En ese marco, se lo ha identificado como un antecedente ineludible del lenguaje humano (Kita, 2003; Tomasello, 2008; McNeill, 2005; Arbib, 2012). Ahora bien, si pudiera mostrarse que al menos en algunas lenguas este tipo de gestos desempeña una función lingüística estructural, entonces no solo tendría relevancia para explicar los antecedentes evolutivos del lenguaje humano o las primeras herramientas para su aprendizaje. En este trabajo, nos interesa examinar un caso contemporáneo de empleo del gesto deíctico en una lengua amazónica, porque muestra que un gesto natural es convencionalizado, a partir de lo cual comienza a formar parte de la gramática de esta lengua. Tomar un caso claro en el que un gesto natural adquiere carácter gramatical combinado con las expresiones verbales en un uso compartido por una comunidad lingüística es relevante porque puede proporcionarnos buenas razones para estudiar el lenguaje desde una perspectiva multimodal que incluya el componente gestual/visual en una relación dialéctica con la modalidad oral/auditiva.

La sistematización y estandarización de gestos en una lengua ha sido mayormente estudiada en las lenguas de señas de comunidades sordas. En esos casos, la modalidad visual es el vehículo de la codificación de los elementos lingüísticos convencionalizados. Más allá de la incorporación semántica de los gestos en las lenguas en general, aquí el desafío es analizar la gramaticalización de acciones corporales visuales de los hablantes de esta lengua verbal, mostrando cómo un gesto se incorpora a la sintaxis y estructura gramatical de la lengua. En general, estas prácticas no son descriptas en términos lingüísticos plenos, o se ponen en conexión con la oralidad en términos semánticos. Esto nos conduce a preguntarnos si es plausible pensar en un modelo modalmente híbrido del lenguaje, no solo para este uso particular de un gesto natural, sino un modelo general que modifique nuestra caracterización del lenguaje, incluyendo como componentes genuinos tanto signos vocales como signos manuales. Este fenómeno ha sido ampliamente estudiado en el caso de la creación de lenguas de señas a partir de gestos espontáneos no convencionales, no así en las lenguas verbales (Perniss, Özyurek y Morgan, 2015).

2. El gesto deíctico como gesto natural

De entre los abordajes evolutivos del lenguaje que hacen foco en los gestos como precursores, el de Tomasello (2008) pone mayor acento en el gesto de señalamiento. En una perspectiva poco común, que considera que la comunicación no lingüística precede y explica la aparición y uso del lenguaje (codificado) y no al revés, rastrea el origen de la comunicación humana hasta los gestos naturales. Por otra parte, para nuestro trabajo, la perspectiva de Toma-

sello se vuelve relevante también porque, desde su punto de partida comunicacional, articula de modo complementario los procesos evolutivos con los histórico-culturales. Para explicar el origen de la infraestructura psicológica subyacente a la comunicación humana (que para Tomasello es claramente cooperativa, lo que la hace exclusiva de la especie), es necesario recurrir a explicaciones en términos biológico-evolutivos. Sin embargo, con este tipo de búsqueda no alcanza. Para explicar el origen de las lenguas convencionales también es necesario conjeturar procesos histórico-culturales mediante los cuales determinadas formas lingüísticas se convencionalizaron en ciertas comunidades de habla, de manera que generaron secuencias de esas formas lingüísticas que se transformaron en construcciones gramaticales. A esto se suma la hipótesis según la cual esas convenciones y construcciones se transmitieron generacionalmente mediante el aprendizaje cultural.

Esta complementación de la historia evolutiva con la historia cultural abre el horizonte para incorporar casos como el deíctico en el nheengatú, un caso proveniente de una cultura contemporánea diversa a la nuestra y en la cual, como esperamos mostrar, los rasgos multimodales del lenguaje se hacen más evidentes.

Como señala Tomasello, sabemos que evolutivamente habría sido imposible saltar de las vocalizaciones o los gestos de los grandes simios a las convenciones lingüísticas arbitrarias, sin pasar por una etapa intermedia de gestos humanos cooperativos no convencionales que portaran significado naturalmente y que pudieran servir como una suerte de cimiento natural para el desarrollo posterior de lenguas codificadas, convencionalizadas. Los gestos constituyen ese lugar para rastrear el origen filogenético del lenguaje. Para esto, se consideran algunos que funcionan como “señales gestuales”, es decir, como comportamientos

comunicativos perceptibles en la modalidad visual. Ahora bien, mientras que las vocalizaciones de los primates, según estudios de Tomasello y colaboradores, son rígidas, innatas y solo sirven para expresar, pero no tienen intención comunicativa, la comunicación gestual ya presente en los simios, en cambio, comparte con la humana aspectos fundamentales de su funcionamiento, tales como la intencionalidad y el uso flexible de señales comunicativas aprendidas. De este modo, según Tomasello y colaboradores, a la hora de explicar los orígenes filogenéticos del lenguaje vocal humano deberíamos recurrir a los gestos, concentrándonos en la modalidad visual más que en la modalidad vocal.

A diferencia de otras hipótesis filogenéticas y ontogenéticas, como las de McNeill (2012) y Arbib (2012), que sugieren un itinerario evolutivo que iría desde la pantomima y la gesticulación hasta llegar a un protolenguaje de señas, las investigaciones de Tomasello (2008) apuntan a los gestos, que no acompañan ni sustituyen expresiones verbales, sino que son actos comunicativos plenos por sí mismos. Especialmente en dos tipos que usamos para referir a algo, y que evolutivamente alcanzan formas complejas de comunicación específicamente humanas: deícticos e icónicos. Estos dos tipos aparecen claramente diferenciados en los trabajos de Tomasello.

Los deícticos son aquellos gestos que hacemos para guiar en el espacio la atención de un receptor, hacia algo ubicado en el entorno perceptual inmediato (Kita, 2003; Tomasello, 2008). Los gestos icónicos, en cambio, son las acciones que buscan orientar la imaginación de un receptor hacia algo que, por lo general, no está en el entorno perceptual inmediato, lo que se consigue con ciertos comportamientos, tales como simular una acción, una relación o la forma de un objeto. Es decir, representan cosas reales,

pero en su ausencia. Ambos se consideran gestos naturales, son simples y espontáneos y, aunque muy básicos, son de uso frecuente y comunican con gran éxito.

3. De los gestos naturales a las lenguas convencionales

Los gestos naturales poseen gran potencialidad comunicativa. Tanto los deícticos como los icónicos buscan orientar la atención del receptor o su imaginación hacia algo determinado, constituyen actos de referencia que se proponen conseguir que el receptor infiera la intención social del comunicador, es decir, lo que este quiere que el receptor haga, sepa o sienta. Es por ello por lo que pueden considerarse las formas básicas más rudimentarias de la comunicación lingüística humana. Aparecen tempranamente y configuran las formas primitivas de comunicación tanto en primates como en niños prelingüísticos.

La cuestión que se plantea, entonces, es cómo se pasa de modos de comunicación “naturales”, simples y espontáneos a los modos “convencionales” que se estandarizan y sistematizan en una gramática propia de cada lengua.

Según Tomasello, la comunicación convencional solo es posible cuando los participantes tienen previamente a su disposición gestos naturales, pero, sobre todo, cuando existe una infraestructura de intencionalidad compartida vinculada con ellos, y habilidades para el aprendizaje cultural y la imitación que les permitan crear y transmitir convenciones comunicativas que se comprendan colectivamente. La infraestructura de intencionalidad compartida (que Searle denominó *“we intentionality”*) es considerada por Tomasello la estructura psicológica que subyace a la cooperación; refiere a la condición para participar en algunas formas de colaboración humana “que implican un

sujeto plural, un ‘nosotros’: intenciones conjuntas, conocimientos mutuamente compartidos, creencias también compartidas, que ocurren dentro de un contexto de móviles cooperativos diversos” (Tomasello, 2008, p. 17). Puede tratarse de acciones más o menos simples, desde, por ejemplo, proponerse construir juntos una herramienta o dar juntos una caminata, hasta la comunicación humana que es netamente cooperativa.

La intencionalidad compartida estaría en la base, entonces, de toda comunicación, incluso de la convencional. De este modo, la comunicación se definiría como una actividad conjunta en la cual el comunicador se esfuerza en brindar información del modo más claro posible sobre el referente; al mismo tiempo, el receptor intenta reconocer la intención social del comunicador. En el mismo sentido, cuando consideramos los móviles comunicativos más básicos, advertimos que, en los gestos naturales o las lenguas codificadas, los móviles comunicativos son los mismos; desde este punto de vista, no habría una diferencia cualitativa. En ambos casos, los móviles son pedir, informar y compartir.

Siguiendo la caracterización de Tomasello, “pedir” expresa la intención de que otro haga algo para ayudar a quien inicia la comunicación o de que brinde la información que necesita; “informar” conlleva la intención de que el otro sepa que se cree que la información le ayudará o le interesará; “compartir”, en cambio, manifiesta la intención de que el otro sienta algo de modo que los dos puedan experimentar sentimientos en común. El hecho de considerar que estos fines básicos son comunes tanto a la comunicación gestual natural como a las lenguas codificadas apoya la tesis de que las convenciones lingüísticas arbitrarias solo habrían podido surgir en el curso de la evolución en el contexto de actividades de colaboración, en las

cuales los participantes ya compartieran las intenciones y la atención, actividades que estaban coordinadas por formas naturales de comunicación gestual.

En este esquema comunicacional, el señalamiento y la mimética humanas sucedieron a los gestos de los simios cuando surgió la cooperación considerada exclusiva de la especie humana. El acto de señalar tiene su origen en la tendencia natural de los seres humanos a seguir la dirección de la mirada ajena hacia objetos externos, y la mimética está vinculada con la tendencia a interpretar que las acciones de otros son intencionales. Así, por ser naturales, los gestos son considerados intermediarios entre la comunicación de los simios y las convenciones lingüísticas humanas. Es plausible pensar que los lenguajes convencionales (primero de señas y luego vocales) se fueron montando sobre los gestos que ya se comprendían y luego reemplazaron el señalamiento y la mimética por una historia de aprendizaje social compartido. Ese proceso fue posible gracias a las aptitudes humanas para el aprendizaje cultural y la imitación, que habrían permitido a los individuos aprender de otros y reconocer sus estados intencionales. En algún tramo de esta trayectoria evolutiva, los seres humanos comenzaron a crear y transmitir, a través de la cultura, diversas convenciones gramaticales, organizadas en complejas construcciones, que codificaban expresiones lingüísticas utilizables de modo repetible en situaciones comunicativas.

Esto nos plantea una cuestión. Los gestos naturales deícticos, que caracterizamos como espontáneos y simples, contienen mínima información en sí mismos, comparados con los elementos de los lenguajes humanos convencionales, mucha menos información que los gestos icónicos, e incluso que los elementos de las lenguas de señas. Parecen, por ello, dispositivos de comunicación muy ende-

bles. Cabe preguntarse, entonces, cómo pueden llegar a ser plenamente comunicativos, o, de modo más específico, cómo es que algo tan simple como un dedo extendido puede comunicar información tan compleja y, además, hacerlo de manera tan distinta en ocasiones diferentes.

Como ya mencionamos, la comunicación que establece este gesto tan simple dependerá fundamentalmente del terreno conceptual común y del marco atencional conjunto del comunicador y el receptor en ese momento. Tanto los gestos espontáneos como la comunicación gramaticalizada dependen de la misma infraestructura de intencionalidad compartida. Pero, además de ese *background* psicológico evolutivo, en algunos casos, puede desarrollarse una justificación de orden cultural. Para mostrar un caso en el que el deíctico adquiere un uso convencionalizado particular en una comunidad de habla amazónica, abordaremos el caso de la lengua nheengatú, perteneciente a las tupí-guaraní.

4. Multimodalidad y gramaticalización del deíctico: el caso nheengatú

En este apartado analizaremos un caso de un uso particular del deíctico en la lengua nheengatú, estudiado en el artículo de Floyd “Modality hybrid grammar? Celestial pointing for time-of-day reference in Nheengatú” (2016). Esta investigación ofrece un estudio de caso de una práctica corporal de comunicación en la modalidad visual, usada por hablantes de la lengua indígena brasiliense nheengatú, una de las lenguas de la familia tupí-guaraní, para hacer referencias de tiempo integrando el gesto con el lenguaje verbal. Los hablantes de esa lengua señalan “la posición

del sol” para indicar un momento del día y recorren con el gesto deíctico una trayectoria del arco solar imaginario para señalar duración del tiempo.

Según Floyd, en el uso del gesto de esta comunidad se puede observar que las referencias temporales de esta lengua se basan en el carácter indexical de la acción de señalar, tipo fundamental de gesto humano que cumple la función de acto comunicativo completo y que está presente bajo diferentes formas en todas las sociedades humanas. Floyd aplica a este caso la descripción del gesto de señalamiento que Kita (2003) ha propuesto. Según esta, la función básica del señalar es interpersonal, ya que este gesto tiene la particular intención de orientar la atención del otro hacia algo. Es posible inferir la intención social del gesto. Para ello es necesario averiguar por qué se ha realizado ese acto referencial y qué pretende el comunicador del receptor. Por lo tanto, cada forma de este gesto incluye dos aspectos: una intención referencial (prestar atención a un objeto del entorno), más una intención social. Floyd (2016) y Kita (2003) identifican los gestos deícticos como aquellos movimientos comunicativos que orientan la atención del receptor en el espacio (intención social) y la centran en algún lugar del entorno perceptual inmediato (intención referencial).

Para analizar este uso del gesto de señalamiento para hacer referencias temporales conviene partir de un ejemplo. Floyd considera el relato de un cuento tradicional por parte de un hablante nativo y la traducción que realiza un intérprete del nheengatú al portugués para el investigador. El hablante relata un cuento en el que unos niños son transformados en aves durante la noche y luego son observados por hombres que llegan al día siguiente. En el ejemplo, el verbo “llegan” es modificado por la frase adverbial “en la mañana cuando el sol está aquí”. Alineado con

la palabra “aquí”, el hablante apunta (y alinea su mirada también) al oeste a 65 grados de altitud, aproximadamente, al lugar donde el sol debería estar a las 11 a. m.

En este relato, aparece el señalamiento manual funcionando claramente como una acción comunicativa plena. La descripción de este momento se vuelve doblemente llamativa cuando el ayudante del investigador traduce la hora como “11 en punto” (en portugués). Cuando el investigador le pregunta al ayudante sobre cómo supo la hora, dado que el nativo no la había mencionado, el traductor le contesta que el hablante “dijo” la hora señalando esa posición en el cielo. En ese sentido, el traductor asume el significado de la frase combinando tanto los elementos verbales como el gesto que señala la posición habitual del sol a esa hora. Esta traducción constituye una muestra clara de la comprensión del significado de la expresión en clave multimodal.

En este uso del gesto de señalar, emergen dos aspectos que me interesa analizar. Por un lado, aparece un ejemplo claro de multimodalidad lingüística. El gesto de modalidad visual es convencionalizado y combinado con las expresiones verbales, de modalidad auditiva, como un elemento más de la oración. Las oraciones que marcan momentos o períodos temporales combinan palabras con el gesto de señalar. Para señalar un momento preciso o cierto período de tiempo, no se utilizan palabras, sino que el brazo se alza hacia el lugar que el sol ocuparía en el arco celeste en el momento que se quiere indicar. Esta combinación de gesto y palabra es rápidamente traducida por el intérprete como una hora puntual y reconocida como un enunciado. Es importante mencionar que esta lengua no tiene *mark tenses*, la marca explícita de los tiempos verbales. Las referencias explícitas de tiempo son realizadas a través de modificadores adverbiales y no de marcas de tiempo verbales.

Los hablantes de la comunidad nheengatú complementan palabras que funcionan como adverbios mediante un movimiento corporal de modalidad visual.

Por otro lado, este uso del señalamiento saca a la luz una de las vías por las cuales un gesto natural, espontáneo, básico, es convencionalizado en un proceso que adjudica al gesto una función gramatical estable y estandarizada. Este punto es particularmente importante para analizar uno de los procesos culturales –no solo evolutivos– por los cuales los gestos adquieren carácter plenamente lingüístico.

Según el autor, la grammaticalización del gesto necesita como paso previo la convencionalización de un gesto en origen natural. Esta grammaticalización no sería explicable si no tenemos en cuenta que esta lengua es multimodal.

La grammaticalidad de este uso del señalamiento es el resultado de la estandarización del gesto, logrado gracias a la multimodalidad en la combinación de elementos visuales y auditivos y la emergencia de un único significado entre ambas modalidades –aspecto directamente vinculado con el anterior–.

Veamos cada uno de estos indicios de grammaticalidad:

- Estandarización del gesto: todos los miembros de la comunidad llevan a cabo el gesto del mismo modo, respetando los mismos ángulos, sin importar el lugar donde se encuentren, ni la visibilidad del sol.
- Multimodalidad visual-auditiva: los hablantes nativos combinan la modalidad manual-visual del gesto con una frase de modalidad vocal-auditiva.
- Un único significado emergente: de esa combinación, no resulta una yuxtaposición o redundancia de significados, sino un único contenido semántico. La unidad resultante condensa el significado lingüístico en esta combinación de dos modalidades, auditiva y visual.

Aquí nos parece de gran potencia explicativa la teoría de McNeill (2005) que propuso el concepto de *growth point*. Según este autor, el gesto es un componente que integra el lenguaje y no un mero acompañante. El gesto se sincroniza con el habla y se vuelven coexpresivos. El lenguaje mismo es definido en la dialéctica entre imagen y habla (*image-speech*). Ambos componentes, gestuales y verbales, constituyen modos semióticamente diferentes y hasta opuestos. Mientras los gestos están ligados a la imagen, son globales, sintéticos, aditivos, instantáneos; el habla es composicional, analítica, combinatoria y lineal. Este carácter dual genera una tensión dialéctica entre los dos modos, una combinación inestable que busca resolución. Esta tensión impulsa el pensamiento al acto de habla, a la formulación de una expresión lingüística. La sincronización entre ambos establece una combinación de modos opuestos que se intersectan en un punto (*growth point*) en el que expresan la misma unidad-idea. Es decir, no tenemos lenguaje verbal, expresión lingüística sin un componente del modo gestual/imagen. Ambas dimensiones son necesarias para el estudio de esta lengua. La unidad dialéctica inestable que generan ambos componentes muestra la dimensión dinámica del lenguaje, enfocado multimodalmente, como un proceso en contexto y en tiempo real. Los gestos proveen las imágenes y la tensión dialéctica misma impulsa el ulterior decurso del habla y el pensamiento. Esa dialéctica muestra una sincronización sistemática en puntos de máxima fuerza discursiva al momento de hablar, donde ambas modalidades expresan significados compartidos subyacentes pero no redundantes, y así se genera un mayor dinamismo comunicativo. En este caso, aunque ocurren simultáneamente, el significado aportado por la modalidad visual no está presente en la modalidad auditiva.

Otro ejemplo ilustrativo con el que se identifica cómo opera el gesto de señalamiento temporal de la comunidad combinado con emisiones lingüísticas, y cómo es incorporado en un registro gramatical, es el relato de la fabricación de harina de mandioca que dura tres días. En el relato reconstruido por Floyd (2016), pueden aparecer simultáneamente con indicaciones temporales visuales tres tipos de elementos verbales:

1. Vocalizaciones deícticas que requieren referentes indexicales provistos por el señalamiento al cielo: “Cuando el sol está allí” (significando “A las 11 a. m.”).
2. Adverbios de tiempo generales cuyos significados son precisados por gestos referenciales de tiempo de grano fino: “Por la mañana”.
3. Predicados en los cuales la referencia temporal visual es el único modificador que ocurre simultáneamente sin el apoyo de otras emisiones de habla. En esos casos, los modificadores adverbiales visuales, generalmente, preceden al verbo. En otros casos, la emisión lingüística incluye elementos adverbiales no tan precisos, como los señalados en el punto anterior, y el señalamiento hacia el cielo ocurre simultáneamente, ocupando la misma posición sintáctica como modificadores en la emisión. Cuando no se emiten oralmente elementos adverbiales, esto permite expresar la referencia temporal simultáneamente con el predicho, una *affordance* de la modalidad visual imposible para los elementos verbales: “Se termina el proceso” (mientras se señala el lugar del sol al atardecer).

En la breve descripción de este relato se hace visible el uso combinado de las modalidades visual y auditiva en el uso gramatical del gesto.

Ahora bien, pensar la gramaticalización del gesto deíctico abre un problema interesante. Tal como presentamos al comienzo, este gesto es de un tipo natural, es decir, básico, evolutivamente primitivo, prelingüístico, simple, no nace de ningún código preestablecido entre los interlocutores, ni lingüístico ni de ningún otro tipo (Kita, 2003; Tomasello, 2008). Comparte estos rasgos con los gestos icónicos, la mímica, que constituyen también un tipo de gesto natural, pero, a diferencia de estos, contiene mucho menos significado en sí mismo ya que solo dirige la atención hacia un objeto disponible en el entorno perceptivo inmediato (Tomasello, 2008). Si bien los gestos naturales deícticos involucran una enorme diversidad y complejidad comunicativa, es ampliamente aceptado que la comunicación humana mediante el lenguaje implica dispositivos convencionales que actúan como medios para coordinar la atención y la acción. En este sentido, la teoría debe dar cuenta del proceso de convencionalización que fija ese movimiento del brazo y del dedo para indicar, de manera estable y regida por normas de uso correcto, un momento del día.

Por otra parte, es preciso explicar cómo puede este gesto adquirir un uso estable en la lengua amazónica que exprese un contenido representacional que no refiera a un objeto disponible perceptualmente, cuando esto último era, precisamente, lo que definía y diferenciaba a los gestos deícticos de los gestos icónicos. Pensamos que el uso gramatical del gesto de señalamiento incorporado en la lengua hace que el deíctico pierda la intención referencial original. Pero, a diferencia de Floyd, que sostiene que el carácter de gesto deíctico se mantiene en ese proceso de gramaticalización, nosotros proponemos que el movimiento que antes señalaba, al adquirir un uso gramatical, deja de ser un gesto deíctico por el hecho de que ya no señala nada accesible del entorno circundante. A

continuación, reconstruiremos y luego discutiremos la vía que ensaya Floyd para explicar este proceso de gramaticalización del gesto natural deíctico.

El autor parte del componente intencional del gesto deíctico. El hecho de que el acto de señalar pueda integrarse de maneras muy variadas y complejas en las diversas situaciones diarias subraya cierta maleabilidad de los gestos deícticos. Floyd supone que esto es posible porque hay una doble intención en el gesto: la intención referencial y la intención social. El comunicador intenta dirigir la atención del receptor hacia algo por alguna razón y el receptor trata de atender al señalamiento e inferir esa razón, a veces saltando una gran distancia referencial. La posibilidad de que el deíctico comunique sin señalar nada directamente disponible perceptualmente se asentaría en el hecho de que se cuente con diversos tipos de conocimientos previos que deben constituir un terreno común compartido entre hablante e intérprete.

Como los gestos de señalamiento implican menos contenido referencial en el signo mismo, para su interpretación dependen del terreno compartido. Asimismo, las lenguas también están plagadas de expresiones, tales como los pronombres, que dependen totalmente de un contexto compartido para su interpretación. Hasta ahora, se había considerado la referencia a entidades desplazadas en el espacio y en el tiempo como un rasgo exclusivo de las lenguas convencionales, dado un contexto compartido adecuado. Sin embargo, las investigaciones de campo realizadas por Floyd (2016) pretenden mostrar que los individuos pueden señalar, además de hacer gestos icónicos, para orientar la atención de otro hacia entidades que no están presentes en ese momento, o que podrían hacer gestos concebidos, incluso, para indicar de manera directa entidades ausentes, como el lugar del sol durante el día.

Ahora bien, no está claro que el gesto deíctico en cuanto tal pueda proporcionar esta evidencia. La cuestión que nos interesa plantear ahora es cómo el gesto deíctico, que por definición dirige la atención a algo accesible en el entorno inmediato, podría usarse para señalar algo ausente. Tratar de aclarar esto requiere desentrañar cómo el señalamiento llega a formar parte de una lengua, y evaluar si, en este proceso, el uso del señalamiento pierde el carácter deíctico, ya que deja de señalar un objeto accesible perceptualmente y adquiere la función semántica de indicador de tiempo.

¿Cuáles serán los criterios lingüísticos que satisface el gesto, en este caso, para llegar a integrar una lengua convencional? Como veremos, desde distintas líneas de investigación se han propuesto diferentes criterios. Floyd repasa algunos, entre los que destaca el de McNeill (2005), quien ofrece un criterio propio para los gestos, diferenciándolos de las construcciones lingüísticas: a) un gesto transmite significado de modo global, mientras que las construcciones lingüísticas componen significados combinando subunidades; b) los gestos individuales no se combinan entre sí, mientras que los elementos del sistema lingüístico se combinan productivamente; c) las relaciones convencionales forma-significado son más estables en los elementos lingüísticos convencionales y más dependientes del contexto en los gestos; d) los elementos del sistema lingüístico tienen estándares de formas, en cambio los gestos son más idiosincrásicos.

Finalmente, Floyd elige el criterio de Okrent (2002) para analizar el caso del deíctico en esta comunidad. De acuerdo con este, la convencionalización lingüística se da en la medida en que los hablantes como grupo muestran estándares de forma, significado uniforme y productividad, independientemente de la modalidad de expresión. Sobre la base de esto, lo que se vuelve crucial evaluar

en el caso del señalamiento de los momentos del día en el nheengatú sería, básicamente, si los hablantes de esta comunidad reconocen formas estandarizadas del gesto que rijan la manera en la que este se realiza, de tal modo que pueda evaluarse cuándo el gesto es equivocado; si todos los hablantes los producen y los comprenden de modo uniforme, es decir, si poseen un significado estable; si generan productividad; y si se incorpora el gesto en diversas combinaciones según las cuales varíe su significado.

Siguiendo los requisitos de Okrent, podemos considerar que sí se justifica tratar el uso temporal del señalamiento como parte de la lengua convencional dado que:

- cada vez que los hablantes de esta lengua se refieren a una hora específica del día, usan una construcción que incluye articulaciones visuales y auditivas simultáneas;
- el elemento auditivo es una frase, que incluye mínimamente un predicado y, opcionalmente, modificadores adverbiales;
- el elemento visual consiste, mínimamente, en apuntar, indicando la posición del sol en el arco este-oeste, a veces indicando dos posiciones y la trayectoria entre ambas;
- elementos de ambas modalidades se combinan compositivamente de acuerdo con restricciones específicas, y el elemento visual agrega un significado adverbial convencional, independiente del contexto y no redundante; y
- si se violan esas restricciones, los hablantes muestran sensibilidad a la desviación de la forma estándar.

Podemos ver aquí que las referencias temporales del nheengatú constituyen un sistema convencionalizado expresado en la modalidad visual cuyas propiedades son más comparables a las de las lenguas de señas que a aquellas usualmente atribuidas a los gestos que acompañan el habla (gesticulación). Este sistema, además, constituye un subsistema del lenguaje verbal, y la única diferencia substancial entre las referencias de tiempo visuales, de grano fino y las referencias verbales, más generales, proviene de sus modalidades de expresión, no de diferencias gramaticales. La referencia temporal del nheengatú equivaldría, básicamente, en los aspectos más relevantes, a elementos de un tipo de lengua de señas especializado.

De acuerdo con la propuesta de Floyd, entonces, con base en los cinco puntos mencionados, podemos argumentar que esta práctica corporal cumpliría, efectivamente, los requisitos de la convencionalización. Esto prueba, por un lado, que, además de las lenguas de señas, hay otros casos, como es este, en los que la capacidad humana para la expresión lingüística puede mostrar un híbrido que conjunta ambas modalidades, la visual y la auditiva. En segundo lugar, se desprende que el sistema de referencias temporales visuales del nheengatú es parte de un sistema lingüístico más amplio de modificadores del predicado, incluyendo adverbios de tiempo orales. Esto representa un gran desafío para la comprensión tradicional de la división entre gramática y uso en el que la contribución de la conducta corporal visual para significar es usualmente pensada como perteneciente al nivel pragmático, no así al morfosintáctico.

5. Incorporación gramatical del gesto deíctico

Para responder la pregunta sobre cómo se produce la incorporación gramatical del deíctico, Floyd elabora una hipótesis acerca de cuál podría haber sido el proceso por el que se logra fijar un significado al gesto deíctico, independizándolo de su referencia a un objeto perceptualmente accesible. El autor aborda la explicación de este proceso sin abandonar el carácter deíctico del gesto. Y recurre a una explicación gradual de la incorporación gramatical del elemento visual. Sin embargo, dado que este deíctico no refiere a nada en el entorno inmediato, su particular función referencial, cuando este se gramaticaliza, tiene que ser explicada de otra forma.

En el artículo de Floyd (2016), se realiza una génesis de cómo el acto de señalar con el brazo ha llegado a adquirir una función de indicador temporal. En el origen de las expresiones lingüísticas multimodales de señalamiento de tiempo que actualmente se observan en la lengua nheengatú, los elementos deícticos visuales comienzan apareciendo de modo dependiente de los elementos deícticos verbales. Es decir, en el origen se acompaña el movimiento que señala el lugar del sol con una palabra deíctica. Esas construcciones de referencias temporales visuales van adquiriendo un carácter cada vez más independiente de la modalidad verbal. Cuando el gesto aparece con palabras que funcionan como adverbios, el señalamiento al cielo no está explícitamente vinculado a una palabra deíctica, pero su co-ocurrencia con un adverbio temporal provee una pista de que el significado deseado es temporal. Cuando una referencia temporal visual modifica un predicado, se da su forma gramaticalizada más independiente del habla. En ese caso, el señalamiento es

producido y comprendido de un modo no ambiguo como una referencia temporal sin necesidad de elementos verbales (Floyd, 2016).

Esta descripción por pasos graduales sobre cómo un gesto natural pudo haber adquirido un carácter convencional plenamente lingüístico debe ser complementada con alguna hipótesis al estilo de la de Tomasello (2008) sobre la habilidad de los seres humanos para crear con otros individuos diversas formas de terreno conceptual común y atención conjunta. Sobre esa base, se asienta la capacidad comunicativa del simple acto de señalar. Ese terreno común sería necesario también para que este gesto deíctico del nheengatú adquiera significado en sí mismo, y, al mismo tiempo, sería la clave para pensar si este y otros tipos de gestos podrían tener funciones gramaticales en otras lenguas convencionales.

A partir del análisis de este caso, Floyd (2016) concluye que el empleo de la modalidad visual para las funciones numéricas abstractas, como medir unidades de tiempo, que en otras lenguas se completa con palabras o términos numéricos, puede mostrar el carácter plenamente lingüístico que pueden adquirir culturalmente algunos usos de los gestos deícticos. Esto ampliaría la evidencia de casos de gramaticalizaciones de gestos, más allá de los sistemas de las lenguas de señas.

En resumen, Floyd desarrolla sus argumentos para sostener que en esta comunidad se opera un proceso efectivo de convencionalización del gesto natural deíctico en cuanto tal, por el cual este se integra en el lenguaje verbal al adquirir un uso convencional y una función gramatical, gracias a una integración multimodal. El autor no cuestiona el carácter deíctico del gesto en el desarrollo de su estudio.

6. Gestos metafóricos

Ahora bien, hemos partido, al inicio de este trabajo, de la distinción planteada por Tomasello entre gestos naturales deícticos e icónicos. Hemos remarcado que el carácter natural de los gestos estriba en que estos son simples y espontáneos, en contraposición a los gestos convencionales, que se estandarizan y sistematizan en la gramática propia de cada lengua; que los gestos naturales tienen carácter prelingüístico; que contienen mínima información en sí mismos, presentándose sumamente dependientes del contexto intersubjetivo compartido, y que los deícticos en particular no contienen en sí prácticamente información, sino que conllevan la intención de dirigir la atención hacia algo accesible perceptualmente en el entorno inmediato, a diferencia de los gestos icónicos, que, a pesar de ser tan básicos como los deícticos, dirigen la atención hacia un aspecto, una circunstancia, una acción que no está presente en el entorno inmediato.

Dada esta caracterización del gesto deíctico, nos preguntamos si la argumentación de Floyd es suficiente para apoyar la tesis de que sigue siendo un gesto deíctico el que resulta convencionalizado e incorporado a la gramática de la lengua nheengatú. A pesar del enfoque gradual de Floyd, mediante el cual intenta mostrar una paulatina independización del gesto de su referente, parece difícil pensar que el gesto de señalamiento siga siendo deíctico cuando ya no está señalando un objeto accesible en el momento en el que es usado. La convencionalización del gesto de señalamiento en la lengua nheengatú sí parece suficientemente demostrada en el estudio de Floyd. Nuestra duda apunta a la permanencia del carácter deíctico de dicho gesto. ¿Seguiría tratándose de un gesto deíctico a pesar de no referir a nada presente en el entorno?

Para elucidar el carácter que adquiere el señalamiento en este proceso por el cual pasa a formar parte de la lengua convencional nheengatú, propongo explorar la posibilidad teórica de que ese uso del señalamiento en la comunidad amazónica alcanzara la convencionalización por otro proceso que no haya tenido en cuenta este autor. Me refiero, precisamente, a evaluar si el señalamiento que se incorpora a la gramática convencional podría haberse convertido en otro tipo de gesto, un gesto metafórico. Dicho desarrollo nos permitiría una mejor explicación del proceso de convencionalización que no dejara de lado el carácter propio de un deíctico de señalar un objeto presente, pero que pudiera dar cuenta de cómo el señalar puede adquirir una función gramatical en la que ya no fuera el carácter deíctico su cualidad principal.

Floyd rechaza la posibilidad de que el gesto de señalamiento sea un gesto metafórico (Floyd, 2016, pp. 36-37), identificando por gesto metafórico el tipo convencional “emblema”, que es arbitrario, porta significado en sí mismo según las convenciones lingüísticas de una comunidad y funciona de modo independiente del habla. Es el caso del puño cerrado con el pulgar arriba en señal de OK que, independientemente de la articulación con las emisiones verbales, tiene un significado preciso y delimitado. Este no parece ser el caso del señalamiento que analizamos, porque este último funciona articulado con el habla. Se trata más bien de explicar la convencionalización de un gesto que pasa a formar parte de la lengua verbal de carácter híbrido.

Acudiendo también a los gestos icónicos metafóricos, propongo aquí argumentar a favor de la posibilidad de que el señalamiento visual temporal nheengatú posea carácter metafórico, pero en cuanto metáfora conceptual en el sentido propuesto por Cienki y Müller (2008). Este modo

de considerar a ciertos gestos como metáforas puede ser visto como una expansión de ciertos signos en la modalidad visual del enfoque sobre las metáforas conceptuales de Lakoff y Johnson (1980), que trata las metáforas conceptuales como una forma de cognición corporizada. En ese orden, también se estudia el gesto con relación al pensamiento.

Ahora bien, ¿a qué nos referimos cuando hablamos de gestos metafóricos? Típicamente, se identificó a los gestos metafóricos con movimientos que representan conceptos abstractos a partir de una modalidad perceptual. Más precisamente, movimientos de las manos que indican o representan el dominio fuente de una metáfora conceptual (McNeill, 2005, p. 154).

Justamente uno de los primeros en reconocer esta función de los gestos fue el mismo Wundt (1900), quien se refirió en particular a gestos que indican conceptos temporales en términos espaciales. Este podría ser el caso para las referencias temporales de la lengua amazónica, en la que se habrían desarrollado conceptos temporales (sobre momentos o períodos de tiempo) en términos espaciales (apuntar con el brazo hacia una posición en el arco solar imaginario). Si tenemos en cuenta que, por definición, el gesto deíctico tiene por intención señalar algo presente, parece razonable pensar que, una vez que se ha perdido el objeto de la intencionalidad referencial, el carácter del gesto ya no habría de ser considerado deíctico, sino metafórico.

La especificidad del gesto metafórico, a pesar de compartir con los gestos icónicos su carácter pictórico (McNeill, 1992), es que representan un dominio abstracto. En el caso de los gestos icónicos, el referente es una acción, entidad o relación concretas, o ciertos aspectos de estas. Ahora bien, de modo más general, podría pensarse

que la metáfora conceptual constituye un procedimiento cognitivo consistente en entender una cosa en términos de otra. Ello implicaría que los gestos metafóricos serían aquellos que incorporan un mapeo multidominio (Cienki y Müller, 2008). Los gestos metafóricos son definidos (Cienki y Müller, 2008) como movimientos voluntarios del cuerpo que realizan un mapeo multidominio para expresar pensamientos abstractos o sentimientos.

Con respecto a la multimodalidad, Cienki y Müller consideran que gesto y palabra pueden presentar relaciones variables con respecto a la metaforicidad. Las metáforas conceptuales gestuales pueden ser semánticamente coexpresivas con las emisiones verbales, aunque esto no obligue a la sincronicidad o simultaneidad de gesto con palabra. Pueden compartir la carga comunicativa para expresar la misma metáfora multimodalmente apareciendo como una expresión gesto-verbal metafórica (Müller, 2004). Sin embargo, Müller también acepta que la metáfora puede estar en el gesto sin necesidad de la co-ocurrencia de un elemento verbal.

Entre las clasificaciones de los tipos de gestos metafóricos que Cienki y Müller analizan, la más apropiada para identificar y analizar el gesto temporal del nheengatú es la de gestos “etimológicos” (Cienki y Müller, 2008). En ese caso, se expresa la metáfora en el movimiento gestual y se combina con las palabras para expresar los conceptos abstractos. Se entiende por gesto etimológico aquel movimiento corporal que expresa el origen concreto de un concepto. Centrarnos en este tipo de gesto metafórico posibilita identificar el movimiento de apuntar hacia lugares precisos del eje este-oeste que representa el arco solar, como una representación del gesto original de señalar el

sol según los momentos del día, es decir, el origen “etimológico” del concepto temporal abstracto que finalmente se quiere expresar, pero ya omitiendo el referente original.

Otro aspecto favorable que justificaría considerar este señalamiento como un gesto metafórico estriba en que, siguiendo a Cienki y Müller, los gestos metafóricos no tienen que ser universales. Pueden ser culturales, tal como la expresión aymará de pasado y futuro, que ubica el gesto que indica futuro hacia la espalda y el de pasado hacia delante, en dirección de la mirada. En el caso de la comunidad amazónica, también parece depender de las bases de la experiencia de esa cultura. Precisamente las referencias espaciales utilizadas para referir al tiempo pueden estar vinculadas con las diferencias geográficas propias del entorno natural de cada comunidad, ya que según el entorno se modifica el marco para las referencias espaciales. En el caso elegido, la comunidad habita una zona cercana a la línea del Ecuador y eso facilita la estabilidad de la referencia al arco solar imaginario.

Por otra parte, un aspecto relevante para nuestro trabajo se centra en la multimodalidad. Anteriormente, vimos la necesidad de estudiar el caso del señalamiento temporal visual desde una perspectiva multimodal. Ese enfoque no se pierde si consideramos este gesto como metafórico antes que como deíctico. Las metáforas conceptuales están relacionadas con un principio cognitivo corporizado general más que con solo una propiedad del lenguaje. Las metáforas gestuales de este estilo están atravesadas por la multimodalidad, ya que, a través de diversos vehículos (palabras, movimientos corporales), representan un concepto abstracto. El contenido pictórico de estos gestos representan una idea abstracta, “una imagen de lo invisible” (McNeill, 1992, p. 14). Las metáforas conceptuales pueden aparecer plasmadas en combinaciones de

modalidades, como es el caso de las metáforas verbales-gestuales (Müller, 2004b) que nos interesan, aunque también puede ser el caso de que las metáforas se den solo en el registro verbal o en el gestual.

La eficacia de los vehículos multimodales es estudiada desde este enfoque de la metáfora conceptual gestual (véase Cienki y Müller, 2008) en estudios experimentales que dan apoyo a la idea de que tales vehículos favorecerían la comunicación, ya que los receptores de las emisiones lingüísticas decodifican la información gestual contenida. Y esta información brindada gestualmente en las comunicaciones cara a cara tendría mayor impacto que la verbal, y sería recordada más fácilmente (esto es desarrollado en estudios empíricos por Beattie, 2003; Beattie y Shovelton, 2001).

Por otro lado, un abordaje multimodal que incorpore el estudio de los gestos en el lenguaje puede abordar la interacción de gestos, palabras y pensamiento configurándose mutuamente y convergiendo en la expresión de la unidad-idea (McNeill, 2005), en la cual interactúan el modo de pensamiento proposicional con el modo imaginístico. Lejos de ser un proceso unidireccional y lineal, esa interacción multimodal cristaliza interrelaciones dinámicas entre pensamiento y habla en la comunicación. En este punto, Cienki y Müller retoman, al igual que Floyd, el concepto de *growth point* de McNeill (1992), que involucra la intersección de dos o más modalidades en un punto significativo que constituye el núcleo de la comunicación.

Esto es relevante porque da la pauta de que puede seguir manteniéndose la potencia de un abordaje multimodal sin comprometerse con la permanencia del gesto deíctico, ya que la potencia comunicativa del señalamiento visual del tiempo, en particular, no se debería a la posibilidad hipotética de señalar el desplazamiento o la ausencia

del objeto -tal como Floyd lo interpreta-, sino a que los gestos describen espacialmente el dominio fuente de las metáforas más claramente que las expresiones verbales con las que se combinan. Algunos de los gestos metafóricos, como es el caso del nheengatú, toman su forma de las actividades corporales diarias y las recrean de manera icónica.

Entre los modos de representación gestual metafórica, Müller (1998) describe el caso de las manos que dibujan rutas en un mapa, o señalan las subas y bajas de una relación de amor como un gráfico. Bien podríamos incluir bajo esa modalidad el caso del señalamiento visual del tiempo. Lo interesante es que estos modos de representación no son solo usados para describir actividades concretas u objetos o propiedades concretas, sino para expresar conceptos metafóricos abstractos, como es nuestro caso de la representación del tiempo. El proceso de base de esta operación está en la hipótesis que afirma que las metáforas conceptuales se generan a partir de la experiencia y la acción corporal (Gibbs y Berg 2002).

7. Conclusiones

Hemos recorrido un camino que comenzó con dos objetivos. Por una parte, hemos intentado argumentar en favor del carácter multimodal del lenguaje mediante el estudio de una lengua en la cual el gesto originalmente deíctico es incorporado a la gramática y adquiere coexpresividad convencionalizada con las emisiones lingüísticas. En el marco de la amplia investigación disponible sobre el origen gestual del lenguaje, nos apoyamos en la distinción elaborada por Tomasello entre gestos naturales y lenguas convencionales. Y, dentro del conjunto de los gestos

naturales, señalamos la distinción entre gestos deícticos e icónicos, consistente en que los deícticos expresan la intención referencial de señalar un objeto disponible en el entorno perceptual inmediato, mientras que los icónicos pueden dirigir la atención a objetos no presentes mediante una representación por semejanza de los objetos o sucesos referidos. Nuestro primer objetivo fue ampliar las vías argumentativas que favorecen un enfoque multimodal del lenguaje. Para ello, analizamos los resultados de la investigación sobre un caso de uso del gesto deíctico propio de la comunidad nheengatú de la Amazonia porque presenta una función particular del señalamiento para expresar momentos puntuales o períodos de tiempo. A partir de ahí, planteamos nuestro segundo objetivo: examinar posibles alternativas sobre el carácter particular del gesto de señalamiento temporal visual en la comunidad nheengatú, y su proceso de convencionalización en la lengua. Pusimos en tensión la posición de Floyd (2016) que sosténía la permanencia del carácter deíctico de este tipo de gesto a pesar del desplazamiento de la referencialidad, con un enfoque que nos llevaría a analizar dicho gesto (Cienki y Müller, 2008) como una metáfora conceptual. A partir de allí, encontramos que ambos enfoques adoptan un abordaje multimodal del lenguaje. Sin embargo, según intentamos mostrar, el enfoque de los gestos metafóricos (Cienki y Müller, 2008) nos proporciona un marco más plausible para considerar el caso de la incorporación del gesto natural deíctico a una lengua convencional. Se trata de un proceso complejo ya que debe dar cuenta de un elemento convencional con base en un gesto natural. La convencionalización de este gesto supondría la transformación del gesto deíctico en metafórico. El movimiento del brazo del hablante no podría considerarse un deíctico cuando no señala nada presente. El esquema de explicación de la

metáfora permite analizarlo en términos de un dominio fuente en el que actúa el movimiento corporal voluntario, el gesto, y de un dominio destino, abstracto, que es el que se representa. Esta proyección, en nuestro caso, presenta la posibilidad de partir de un gesto referencial espontáneo hacia un concepto tan abstracto como puede ser el tiempo. Desde esta perspectiva, el hecho de que el señalamiento pierda su referente no acarrea dificultades, ya que podríamos pensar que se trata de una metáfora gestual de tipo etimológica. Es decir, el movimiento del brazo que se originó para señalar el sol constituye el dominio fuente de la metáfora conceptual para indicar, de una manera más abstracta, el tiempo.

De este modo, partiendo del estudio de un caso de convencionalización del señalamiento en la lengua nheengatú, creemos que la combinación o hibridez entre distintas modalidades incorporadas en la estructura gramatical de la lengua estudiada evidencia la adecuación de un enfoque multimodal del lenguaje. El análisis del señalamiento temporal nos condujo a la identificación de la metafóricidad de algunos gestos, en especial de este tipo de señalamiento temporal visual que parte de un movimiento espacial para representar un contenido abstracto, el tiempo. A la luz de lo expuesto, pensamos que este carácter metafórico conceptual del gesto proveería una clave teórica plausible y potente para el estudio de los procesos de convencionalización de los gestos que se integran en las lenguas.

Referencias bibliográficas

- Arbib, M. (2012). *How the Brain Got Language: The Mirror System Hypothesis*. Nueva York: Oxford University Press.

- Beattie, G. (2003). *Visible Thought: The New Psychology of Body Language*. Londres: Routledge.
- Beattie, G. y H. Shovelton (2001). "An experimental investigation of the role of different types of iconic gesture in communication: a semantic feature approach". *Gesture*, 1, pp. 129-149.
- Cienki, A., y Müller, C. (2008). "Metaphor, Gesture, and Thought". En R. W. Gibbs (Ed.), *The Cambridge Handbook of Metaphor and Thought*, Cambridge Handbook in Psychology, (pp. 483-501). Cambridge: Cambridge University Press.
- Floyd S. (2016). "Modality hybrid grammar? Celestial pointing for time-of-day reference in Nheengatú". *Language* 92, pp. 31-64.
- Gibbs, R., y Berg, E. (2002). "Mental imagery and embodied activity". *Journal of Mental Imagery*, 26, pp. 1-30.
- Goldin-Meadow, S. (2003). *Hearing Gesture: How Our Hands Help Us Think*. Cambridge, MA, US: Belknap Press of Harvard University Press.
- Kendon, A. (1987). "Simultaneous speaking and signing in Warlpiri sign language users". *Multilingua*, 6: pp. 25-68.
- Kendon, A. (2008). "Some reflections on the relationship between 'gesture' and 'sign'". *Gesture*, 8(3), pp. 348-366.
- Kita, S. (2003). "Pointing: A foundational building block of human communication". En Kita, S. (Ed.). *Pointing. Where Language, Culture, and Cognition Meet*. pp. 1-8. Hillsdale NJ: Erlbaum.
- Lakoff, G. y Johnson, M. (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.
- McNeill, David. (1985). "So You Think Gestures are Non-verbal?". *Psychological Review*, 92, pp. 350-371.

- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal About Thought*. Chicago: Univ. of Chicago Press.
- McNeill, D. (2005). *Gesture and Thought*. Londres: Chicago Press.
- McNeill, D. (2012). *How Language Began: Gesture and Speech in Human Evolution*. Cambridge Univ. Press.
- Müller, C. (1998b). "Iconicity and gesture". En S. Santi et al. (Eds.), *Oralité et Gestualité: Communication Multimodale, Interaction* (pp. 321-328). Montreal, París: L'Harmattan.
- Müller, C. (2004). "Forms and uses of the Palm Up Open Hand: A case of a gesture family?". En Müller, C. y Posner, R. (Eds.), *The Semantics and Pragmatics of Everyday Gestures* (pp. 233-256). Berlín: Weidler.
- Müller, C. (2004b). *Metaphors, Dead and Alive, Sleeping and Waking: A Cognitive Approach to Metaphors in Language Use*. Habilitationsschrift, Freie Universität Berlin.
- Okrent, A. (2002). "A modality-free notion of gesture and how it can help us with the morpheme vs. gesture question in sign language linguistics, or at least give us some criteria to work with". En R.P. Meier, D.G. Quinto-Pozos, y K.A. Cormier (Eds.), *Modality and Structure in Signed and Spoken Languages* (pp. 175-198). Cambridge: Cambridge University Press.
- Perniss, P., Özyürek, A. y Morgan, G. (2015). "The Influence of the Visual Modality on Language Structure and Conventionalization: Insights from Sign Language and Gesture". *Topics in Cognitive Science*, 7(1), pp. 2-11.
- Seyfarth, R. M. (1987). "Vocal communication and its relation to language". En B. Smuts, D. L. Cheney, R. Seyfarth, R. Wrangham, y T. Struhsaker (Eds.), *Primate Societies*, (pp. 440-51). Chicago: University of Chicago Press.

- Snowdon, C. T., Brown, C. H. y Petersen, M. R. (1982). *Primate Communication*. Cambridge: Cambridge University Press.
- Tomasello, M. (2008). *Los orígenes de la comunicación humana*. Buenos Aires: Katz Editores, 2013.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., y Moll, H. (2005). "Understanding and sharing intentions: The origins of cultural cognition". *Behavioral and Brain Sciences*, 28(5), pp. 675-691.
- Wundt, W. (1973). *The Language of Gestures*. La Haya: Mouton, 1900.

Esquematicidad mimética en las metáforas multimodales

Contribuciones a la teoría de la metáfora conceptual

CAROLINA MAHLER (CIFFyH – UNC)

1. Introducción

En este trabajo, defiendo que en las metáforas multimodales (Forceville y Urios-Aparisi, 2009), consideradas dentro del marco de la teoría de la metáfora conceptual (TMC) (Lakoff y Johnson, 1980), los esquemas miméticos (Zlatev, 2005), normados primitivamente (Mahler, 2017), pueden dar cuenta de la estructuración de estas metáforas de modo más apropiado y consistente que los tradicionales esquemas de imagen (Hampe, 2005) de la lingüística cognitiva, específicamente como dominios-fuente. A continuación de esta sección introductoria, presentaré, en la sección 2, las características fundamentales de las metáforas conceptuales en general, y de las metáforas multimodales en particular, y me serviré de ejemplos para brindar claridad en la explicación de su estructuración. También me referiré al recorte de la noción de modalidad que aplican Forceville y Urios-Aparisi *et al.* (2009) para las metáforas multimodales, y algunas precisiones al respecto de

la problemática general de los dominios en la TMC. En la tercera sección, me ocuparé específicamente de la noción teórica de dominio como elemento constitutivo necesario de las metáforas en general, y, en particular, dada su emergencia en la temprana infancia, del rol de dominio-fuente (o dominio-origen) en las metáforas multimodales. Examinaré diferentes nociones teóricas utilizadas en la literatura para dar cuenta de la estructura de la metáfora conceptual en general, más precisamente, el pensamiento esquemático y su caracterización. En cuanto a las caracterizaciones más importantes para el pensamiento esquemático, me referiré, en primer lugar, a lo que llamo la tradición en la teoría de la metáfora conceptual en la lingüística cognitiva (autores como Lakoff y Johnson durante la década del 80), especialmente los esquemas de imagen y su proyección en las metáforas; en segundo, a la caracterización de máximo nivel de abstracción para el pensamiento esquemático (Langacker 1987; 2008); y, finalmente, al análisis multinivel de la estructura metafórica en la TMC (Kövecses, 2017). Con respecto a la noción de pensamiento esquemático en la semiótica de la cognición, los esquemas miméticos se presentarán como alternativa a los esquemas de imagen en el rol de dominios-fuente de las metáforas multimodales. La comparación entre ambas nociones de pensamiento esquemático (de imagen y miméticos) en la sección 4 sirve a los propósitos de mostrar cuál noción de esquema presenta un vínculo más directo con la diversidad de la experiencia sensoriomotriz. Teniendo en cuenta tres ejes, a saber, la complejidad de las capacidades cognitivas en juego, la abstracción y la subjetividad en la constitución del pensamiento esquemático, concluyo que el pensamiento mimético es caracterizable como corporizado de manera más consistente que el de imagen. A partir de ello, en la sección 5, concluiré que el alcance de la TMC ampliado a la

perspectiva multimodal, y la estructuración de la metáfora multimodal con esquemas miméticos como fuente, enriquecen la TMC en cuanto robustece su enfoque corporizado. Además, esta perspectiva y este tipo de estructuración eluden un riesgo de circularidad explicativa que conllevaría el estudio exclusivo de metáforas monomodales o verbales, que han sido el objeto de indagación privilegiado en la lingüística cognitiva a lo largo de décadas. Pondré en consideración una dimensión normativa y sociocultural para los esquemas miméticos de los infantes, cuyo comportamiento en relación con estos esquemas no puede ser una mera imitación a ciegas o automática, ni una actividad demasiado sofisticada a nivel cognitivo. Finalmente, las conclusiones se hallan en la sección 6.

2. Metáforas conceptuales y metáforas multimodales

La estructura y procesamiento de la metáfora en los planos cognitivo y lingüístico constituyen un problema teóricamente relevante para la filosofía de la mente y del lenguaje, dadas la ubicuidad y alta frecuencia del uso de las metáforas conceptuales, así como su expeditiva inteligibilidad y variabilidad morfosintáctica en el lenguaje natural. Desde la década de los 80, la teoría de la metáfora conceptual (TMC) de Lakoff y Johnson (1980) cambió drásticamente la forma de abordar la metáfora desde la filosofía, las ciencias cognitivas e incluso la lingüística. Según la TMC, la metáfora no es ni un tropo ni un fenómeno lingüístico, sino un patrón fijo de conceptualización (Lakoff, 1993) que configura el mundo de un modo muy particular: atrapa la idea de que pensamos una cosa (dominio-destino) como otra (dominio-fuente), y que este es un proceso cognitivo habitual y cotidiano, anclado al cuerpo. Esta concepción

tualización subyace a diversas expresiones lingüísticas en la lengua natural, o bien, dicho de otro modo, subyace a manifestaciones en la modalidad verbal, que es donde muy habitualmente los lingüistas cognitivos han enfocado su indagación. Sin pretender explicitar aquí una taxonomía exhaustiva de las metáforas, la clasificación más tradicional de las metáforas conceptuales (MC) incluye: en primer lugar, a las orientacionales, que organizan un sistema global de conceptos con relación a otro gracias a un correlato sistemático entre nuestras emociones y nuestras experiencias sensoriomotrices, p. ej. FELIZ³⁰ ES ARRIBA vs. TRISTE ES ABAJO, habitualmente manifiestas en fraseos como “Eso me *levantó* el ánimo”, “*Caí* en una depresión” (Lakoff y Johnson, 2009, pp. 51); en segundo lugar, a las MC llamadas “ontológicas”, que refieren a eventos o acontecimientos complejos, como entidades o sustancias, p. ej., LA INFLACIÓN ES UNA ENTIDAD se suele manifestar en expresiones tales como “La inflación está bajando nuestro nivel de vida” o “Hay que combatir la inflación”, y LA MENTE es una MÁQUINA (Lakoff y Johnson, 2009, pp. 64-66) es una MC ontológica de gran alcance fuera y dentro de las ciencias. Solo es posible referirse a estos acontecimientos o eventos complejos, o cuantificarlos, adjetivarlos o medirlos (y más), en cuanto se los conceptualiza como entidades. Por último, a las MC estructurales, que hacen uso de una idea claramente delineada y ordenada para configurar otra idea diferente, por ej., EL TIEMPO ES DINERO, que se manifiesta verbalmente en, p. ej., “Me estás haciendo *perder* el tiempo” (Lakoff y Johnson, 2009, pp. 44). Estas MC, de acuerdo con la TMC, configuran gran parte de nuestro pensamiento y de nuestros actos.

³⁰ La mayúscula sostenida se utiliza convencionalmente en la TMC para destacar la naturaleza conceptual, no lingüística, de los dominios que componen la metáfora conceptual.

Siguiendo a Forceville y Urios-Aparisi (2009), habría que tomar seriamente la afirmación -y, pienso, el programa que se sigue de esta afirmación- de que la metáfora no es un fenómeno lingüístico, sino “un modo del pensamiento” (Lakoff, 1993, p. 210). A partir de lo anterior, se despliega la posibilidad de un pensamiento metafórico que se manifieste en todas las modalidades sensoriales, y no únicamente en la modalidad verbal. Una MC es monomodal cuando los dos dominios que la componen se manifiestan en una única y misma modalidad. En la literatura sobre el tema, es notoria la atención preferencial que se le ha otorgado a las metáforas monomodales verbales, en las que los dominios-origen y dominio-destino se manifiestan solo lingüísticamente. Una metáfora monomodal también podría manifestarse exclusivamente en el modo visual, i.e. con signos pictóricos en los dos dominios que la componen, así como en cualquier otra modalidad, a condición de que sea la misma tanto en dominio-origen como en dominio-destino. Ahora bien, si admitimos, como ha afirmado la TMC, que las metáforas tienen expresión en el lenguaje -a diferencia de ser de naturaleza lingüística-, el pensamiento metafórico podría manifestarse en todas las modalidades, tales como la gestualidad o la música, entre otras. El programa de la metáfora multimodal afirma que el pensamiento metafórico bien podría encontrar expresión o manifestación en otros modos que no fueran el lingüístico exclusivamente, y que además habrían de ser diferentes en ambos dominios.

Por supuesto, las nociones “modo” o “modalidad” son controversiales y difíciles de definir, en especial porque tradicionalmente se las ha concebido en relación con los cinco procesos o canales sensoriales de percepción o sentidos. En su lugar, y debido a que dicha noción de modalidad no es adecuada para su aplicación a las metáforas

multimodales (MM), emplearé la utilizada por Forceville y Urios-Aparisi (2009), entre otros autores: modalidad o modo como sistema de signos interpretables, p. ej. signos pictóricos, signos escritos, signos hablados, gestos, sonidos no verbales, música, olor, gusto y tacto. Así, cuando en una metáfora multimodal se concibe una cosa como otra, el dominio-origen y dominio-destino deben manifestarse exclusiva o predominantemente en dos modos o modalidades diferentes. A modo de ejemplo, la metáfora podría contener un dominio-origen que se manifieste exclusivamente como música (o podría coocurrir con otra modalidad como imagen también -aunque no necesariamente-), proyectado a un dominio-destino que se exprese en otra modalidad, por ejemplo, signos escritos (de una lengua). Como característica típica de los dominios que constituyen una metáfora multimodal, es casi imposible convertir o “traducir” exhaustivamente un modo al otro modo. Además, en la MM emergen dimensiones de análisis que cobran importancia -dimensiones habitualmente ausentes en la MC monomodal verbal-: la dimensión del vehículo material (p. ej., papel, celuloide, bits y *bytes*, diseño textil), la del hecho de que típicamente aparezcan o pertenezcan a un género (p. ej. aviso) que a su vez pertenece a un discurso (p. ej. publicitario) proveniente de una comunidad cultural o profesional, y, por último, un particular énfasis en la dimensión corporal. Todas estas dimensiones afectan el uso y la comprensión de las modalidades de las MM. Es muy habitual encontrar este tipo de metáforas en avisos publicitarios, aplicaciones de redes sociales, fotomontajes y viñetas cómicas y políticas, entre otros géneros. Sirva a modo de ejemplo una metáfora multimodal para destacar las dimensiones mencionadas anteriormente. La fig. 1 (Eameo, 2018) ha de entenderse como viñeta cómico-política (subgénero) reproducida como meme (género) en

redes sociales (comunidad cultural) en virtud del vehículo (bits), y en especial en virtud de su multimodalidad (signos pictóricos en formato fotográfico y signos lingüísticos añadidos). La composición de la fotografía, incluida la disposición de las tres personas fotografiadas, y el hecho de que está en blanco y negro agilizan la comprensión del todo, con el texto “Los tres chiflados”, como una viñeta de humor político.

Figura 1: Metáfora multimodal con signos pictóricos y signos lingüísticos



En cuanto a la estructura de las metáforas, la noción de dominio es central. Lo que hace que un concepto sea metafórico, estructuralmente hablando, es la proyección (mapping) de un dominio-fuente a un dominio-destino, ambos coexistiendo como elementos constitutivos de una sola estructura conceptual: la metáfora. Por convención, la explicitación verbal de los dominios de una metáfora se realiza en términos de A ES B, con letras mayúsculas, para destacar su carácter cognitivo (i.e., no lingüístico). Cabe

destacar que esto no significa de ningún modo que se trate de un contenido proposicional o que una metáfora pueda evaluarse en términos de condiciones de verdad –al menos no por su estructura conformada por dos dominios-. A modo de ejemplo, en la MC orientacional INCONSCIENTE ES ABAJO, el dominio-destino INCONSCIENTE se hace inteligible en virtud de las contribuciones realizadas por el dominio-fuente ABAJO. Esta proyección unidireccional del dominio-fuente al dominio-destino (Hipótesis de la Invariancia³¹) es lo que le da forma a mucho de lo que percibimos, pensamos y hacemos, y es lo que subyace a ciertas expresiones habituales correspondientes en la lengua natural, tales como “Está bajo hipnosis” (Lakoff y Johnson, 2009, p. 51). Esta, además, es una instancia de una metáfora monomodal, debido a que los dominios que la componen se manifiestan en la misma modalidad, es decir, en formato lingüístico, tanto en la fuente como en el destino.

Ahora bien, una metáfora fundamentalmente implica una “incongruencia”³² entre los dos dominios que la constituyen, fuente y destino, a pesar de la apariencia de isomorfía o analogía entre los dos dominios. El uso sistemati-

³¹ De acuerdo con la Hipótesis de la Invariancia de Turner (1990), también en Lakoff y Turner (1989) y Lakoff (1990), las relaciones biunívocas de carácter metafórico preservan la topología cognitiva (es decir, la estructura esquemática) del dominio-fuente, de modo tal que los componentes semánticos del dominio-fuente siguen siendo coherentes en el dominio-destino. En este sentido, la metáfora conceptual contiene una estructuración isomórfica de ambos dominios.

³² Salvando las distancias teóricas entre la TMC y la perspectiva de innovación metafórica de Paul Ricoeur, la noción de “transición de la incongruencia literal a la congruencia metafórica entre dos campos semánticos” aparece en Ricoeur (1978, p. 147). Nótese que en el caso de metáforas multimodales, y en general para toda metáfora conceptual también, entiendo “incongruencia” como el vínculo arbitrario entre un dominio y otro tomados por separado, que no poseen entre sí ni una relación de identidad, ni de simetría, ni de transitividad. Esta incongruencia se “resuelve” en la proyección metafórica de un dominio a otro, constituyendo un concepto distinto, es decir, la metáfora. La incongruencia entre dominios se manifiesta mediante la imposibilidad de comprensión literal de las MC verbales.

zado y habitual refrenda solapadamente esa aparente isomorfía tanto del concepto metafórico como de las expresiones lingüísticas que frecuentemente lo acompañan. A modo de ejemplo, puede parecer inevitable concebir isomórficamente la escalaridad de las notas musicales y la verticalidad en el espacio, a partir de lo cual hay tonos altos (se sube) o bajos (se baja). La metáfora conceptual puede explicitarse de este modo: LAS RELACIONES ENTRE TONOS SON RELACIONES EN EL ESPACIO FÍSICO. Sin embargo, no hay nada de natural en esta relación entre tonalidad y verticalidad. Según Zbikowski (2009, p. 361), la proyección de un dominio a otro es “arbitraria. Por ejemplo, cuando interpretas notas ‘más altas’ en el piano, mueves la mano hacia la derecha; cuando interpretas notas ‘más altas’ en el violonchelo, mueves la mano hacia el suelo”. Pero además destaco el carácter metafórico de esta manera de conceptualizar las relaciones entre notas musicales, puesto que en realidad no hay nada vertical en esta relación en la música: se trata de una “incongruencia” que, como suele suceder, las MC resuelven, y de hecho posibilita una manera de comprender, en este caso, cómo se relacionan los tonos musicales entre sí. Un ejemplo de metáfora multimodal que utiliza Zbikowski (2009, pp. 360-362) puede oírse en el pasaje coral del Credo de Palestrina (Orchard Music y 1 Music Rights Societies, 2012) desde 2:05 a 2:28. El texto de la partitura incluye la expresión “Y para nuestra salvación, *bajó* del cielo”; desde el comienzo de la enunciación de la expresión lingüística “*descendit*”, las seis voces que componen este pasaje comienzan un descenso de escala (Zbikowski, 2009, p. 360), y culminan como un proceso estructurado dinámica y ordenadamente después de ocho notas, sin cambiar de dirección. Zbikowski interpreta los aportes de cada modalidad: por un lado, la expresión lingüística contribuye a acceder al

dominio orientacional de espacio físico (VERTICALIDAD), pero además al espacio evaluativo asociado al descenso controlado por la metáfora conceptual EL SER ES ORIENTACIÓN EN ESPACIO VERTICAL. De modo sistemático (Lakoff y Johnson, 1980, pp. 43-49), es posible reconocer esta configuración conceptual en la jerarquización ascendente o descendente de los niveles ontológicos en la filosofía platónica, o en la teoría de la emanación neoplatónica, por ejemplo, en coherencia también con las correspondientes apropiaciones de diferentes religiones, p. ej. PUREZA ES ARRIBA, y PECADO ES ABAJO. Por otro lado, los materiales sonoros de la modalidad musical constituyen por sí mismos el significado de descenso gradual y dinámico, significado que la sola palabra "*descendit*" simplemente no puede captar en ese sentido de gradualidad y dinamicidad que aportan los materiales sonoros -o al menos no puede hacerlo de forma corporizada, como lo hace el sonido-. Más allá del aporte de una y otra modalidad, nótese la incongruencia inherente entre los dominios, en cuanto que los sonidos reales, por ejemplo, al bajar escaleras en una casa o al bajar de una montaña, no son ni análogos ni parecidos a los de descenso en la escala tonal.

3. Posibles caracterizaciones de “dominios”: el pensamiento esquemático

Si bien se acepta que las metáforas conceptuales requieren dos dominios en su constitución, no existe consenso sobre una definición nítida de la noción de dominio. En la lingüística cognitiva, con frecuencia se utiliza la caracterización de área de conocimiento afín, basado en y asociado con la experiencia corporizada. Lakoff y Johnson (1980) definen los dominios como *Gestalts* experienciales,

contextualizados en la experiencia humana recurrente. Es notable que, bajo el paradigma lakoffiano de la TMC, los dominios se definan solo en relación con la comprensión: “La comprensión ocurre en términos de dominios de la experiencia completos y no en términos de conceptos aislados” (Lakoff y Johnson, 2009, pp. 117-118). Es decir, los dominios parecen preceder a la comprensión, y contribuyen a que esta ocurra. Así, la función de los dominios es relacional con la comprensión en cuanto la hacen posible. Debido a que las metáforas conceptuales están ancladas a la experiencia básica, corporizada y prelingüística, se requiere que al menos su dominio-fuente sea una estructura rudimentariamente cognitiva, que suponga una habilidad corporizada o relacionada a la percepción de fenómenos. Si es el caso de que al menos algunos dominios son previos a la comprensión, y previos e independientes del lenguaje, entonces deben ocupar un rol primordial en el desarrollo humano y emerger muy tempranamente en la ontogenia. Y, por ello, deberían operar como precursores de un pensamiento más sofisticado. En cuanto al estudio de las MC, el dominio-fuente debería permitir, junto con un adecuado desarrollo cognitivo, la comprensión o la organización global conceptual del dominio-destino (p. ej. en conceptos metafóricos) en un momento posterior de la ontogenia. Las MC cuya comprensión es más directa son las que contienen dominios espaciales simples como ARRIBA, es decir, las metáforas conceptuales orientacionales (Lakoff y Johnson, 1980, p. 56).

La lingüística cognitiva está comprometida con la posición filosófica general denominada “realismo experiencial” (Lakoff y Johnson, 1999), en oposición a una referencialidad objetiva del mundo. Intenta describir cómo los humanos realmente pensamos y organizamos, mediante la experiencia corporizada, nuestro conocimiento, que no

siempre es estrictamente lógico o congruente, ni siquiera en la adultez. Tomando como especial punto de apoyo los principios psicológicos de categorización mediante prototipos de Rosch (1978) para la formación de categorías que pueden hallarse en una cultura, la TMC introdujo la noción de esquematicidad para dar cuenta de la estructura de las metáforas conceptuales. Esta idea de esquematicidad puede rastrearse hasta Kant (1968), en cuanto su función es conectar la sensación y el concepto. Sin embargo, cabe aclarar que la TMC rechaza las posturas cartesianas o kantianas acerca de la mente, al defender una posición corporizada, antidualista, respecto de la cognición. El candidato canónico de dominio-fuente en la TMC es el esquema de imagen (Hampe, 2005; Lakoff y Johnson, 1980) o EI. Los EI son guestálticos (no atómicos), corporizados, preconceptuales, involuntarios, inconscientes y experiencialmente significativos. Son básicos, en el sentido de que capturan los contornos estructurales de la propiocepción y de que están basados en la experiencia sensoriomotriz; p. ej. ARRIBA, CONTENEDOR. En la TMC, los EI son muy recurrentes en la experiencia corporizada cotidiana (Lakoff, 1987). Además, muchos autores piensan que los EI dan lugar y brindan estructura a dominios más abstractos y complejos de la experiencia (Hampe, 2005).

Figura 2: Metáforas conceptuales con esquemas de imagen ARRIBA y ABAJO como dominios-fuente



La caracterización canónica del pensamiento esquemático de imagen es controvertida dentro de la lingüística cognitiva, aún más para caracterizar los dominios-fuente, lo que se evidencia en el hecho de que diferentes autores no utilicen la noción de esquematicidad de manera consistente dentro del enfoque de la lingüística cognitiva para referir a la experiencia más básica o corporizada (Hampe, 2005). Dentro de la lingüística cognitiva, Langacker (1987) se enfoca en la gramática cognitiva, y no en la TMC; sin embargo, define el pensamiento esquemático como “el proceso de extracción de lo común inherente en múltiples experiencias, para arribar a la concepción que representa un nivel más alto de abstracción” (Langacker, 2008, p. 17). Según este autor, la esquematización se da en diferentes grados de abstracción, de modo tal que una estructura cognitiva es más o menos esquemática de acuerdo con el nivel de abstracción logrado por un sujeto a partir de usos puntuales de esa estructura. Así, la forma y significado convencionales de un esquema (como

estructura) son menos específicos, es decir, más abstractos, que los eventos de uso. Por ejemplo, cuando efectivamente nos referimos a un anillo particular, se trata de un uso muy específico del esquema ANILLO. En virtud de las experiencias con múltiples anillos particulares, el sujeto logra abstraer algunas categorías como “anillo de bodas”, o “anillo de gimnasia”, hasta lograr reunir todos los rasgos comunes de ANILLO como esquema (de máxima abstracción), entendido meramente como “entidad circular” u “objeto circular”. Este autor opina que la simpleza del esquema no se basa en la experiencia corpórea, y que los esquemas tampoco son directamente significativos –como aseguraban Lakoff y Johnson (Lakoff y Johnson, 1980; Lakoff, 1987) al respecto de los EI. Antes bien, son de elevado grado de abstracción relativa a conceptos o una categoría, proveyendo menos información que cada uno de los miembros de la categoría, y por ende son compatibles con todos estos últimos (Cienki, 2013). De forma significativa, tanto Langacker como Lakoff y Johnson y la totalidad de autores dentro de la lingüística cognitiva afirmarían que la experiencia es el punto de partida para el pensamiento esquemático. La diferencia estriba, según mi punto de vista, en que Langacker estaría dispuesto a adoptar un enfoque más subjetivo del *construal* del pensamiento esquemático, mientras que la tradición de la TMC seguiría la perspectiva prototípica de Rosch, no tanto para un sujeto, sino para una cultura en un momento dado (Rosch, 1978, p. 29), lo cual deja sin explicar la cuestión ontogenética para la TMC.

Kövecses (2017), por su parte, aborda la cuestión de la escasa claridad terminológica e indefinición conceptual para los elementos que constituyen la MC en la lingüística cognitiva. Su propuesta clarificadora consiste en la diferenciación de niveles de esquematicidad como niveles de análisis de la estructura de las MC. Estos diferentes niveles

funcionarían simultáneamente y sin una rígida delimitación, conformando un *continuum* de mayor a menor esquematicidad, y de menor a mayor especificidad para cada uno de los dos elementos constitutivos de las metáforas conceptuales (fuente y destino). Así, Kövecses propone la siguiente secuencia: *esquema de imagen*, como el nivel más esquemático, seguido por *dominio*, *marco* y, finalmente, *espacio mental*, como el menos esquemático en el *continuum*. La especificidad se daría de manera invertida en este análisis multinivel, del cual resultaría que los espacios mentales son los más específicos, y los EI, los menos específicos. Sin embargo, Kövecses no se detiene en el problema de la abstracción. De hecho, parece asumir, como es habitual en la literatura de la TMC, que los EI resuelven la comprensión de la estructura más abstracta hacia la cual se proyectan (Kövecses, 2017, pp. 7-8). En ese sentido, los EI serían menos abstractos que cualquier otra estructura cognitiva, y, por necesidad, toda otra forma de conceptualización sería más abstracta que los EI; por lo tanto, la estructura sobre la cual se proyecta el EI en el rol de *fuente* de la metáfora conceptual, es decir, lo que habitualmente se denomina el “dominio-destino”, dependería de la previa conformación de los EI para ser comprendidos mediante la nueva estructura, que es la metáfora conceptual total, y, además, ese dominio-destino sería necesariamente más abstracto que el EI. Sin embargo, esta posición es radicalmente opuesta a la caracterización del pensamiento esquemático en Langacker, en la que los esquemas se definen como la estructura cognitiva de máxima abstracción en relación con un uso o referencia específica. Pienso que la clave para comprender lo que separa una caracterización del pensamiento esquemático como la de Langacker (de máxima abstracción) de las de Kövecses y la tradicional de la TMC (de mínima abstracción) es que,

si bien se refieren al mismo tipo de estructura cognitiva, el primero configura su caracterización alrededor de la constitución subjetiva de una estructura cognitiva representacional, mientras que los segundos, alrededor de la constitución cultural sincrónica de una estructura cognitiva que contrapone la característica “abstracción” a (meramente) “percibido con el cuerpo”. Estas divergencias en la caracterización del pensamiento esquemático han suscitado una propuesta alternativa, que se describe a continuación.

Más recientemente, en el marco de la semiótica de la cognición, Zlatev (2002, 2005, 2007a, 2013) ha propuesto la noción de esquema mimético, o EM, como representaciones dinámicas y preverbales que involucran la imagen corporal, p. ej. SEÑALAR, SALUDAR, ALCANZAR (ver fig. 3). A diferencia de los esquemas de imagen, los esquemas miméticos son representaciones de mimesis corporal inducidas por las recurrencias de actos corporales socio-culturalmente relevantes. “Compartidos prerreflexivamente en una comunidad” (Zlatev, 2005, p. 334), los EM surgen del intercambio interpersonal mediante la exterocepción y propiocepción, a medida que el infante va percibiendo recurrencias significativas de estos actos, hasta que finalmente se internalizan. Constituyen un fenómeno fundado muy básicamente en la imitación neonatal en un primer estadio, hasta que, alrededor de los 14 meses de edad, el infante llega a triangular con el adulto y un objeto externo, ejerciendo un cierto control volicional que le brinda mayor precisión y flexibilidad de imitación (Zlatev, 2013). Los EM son dinámicos, en el sentido de que incluyen movimiento o perspectiva del entorno, y juegan un rol importante en la adquisición del lenguaje, u ontogénesis, y en la evolución del lenguaje, o filogénesis, así como en la relación entre experiencia fenoménica y significado compartido con otros; además, también articulan el significado

lingüístico, el comportamiento significativo con el cuerpo y el entorno, todo lo cual permite anclar el significado lingüístico. De hecho, el pensamiento mimético se postula como precursor del lenguaje, aunque también puede coocurrir con el lenguaje. Los EM constituyen una alternativa a los EI, los cuales, según algunas interpretaciones, o bien tienden más hacia la abstracción que hacia la corporeidad (Cienki, 2013), o bien, según mi interpretación, se los puede entender como constructos cognitivos culturales sincrónicos que, en consecuencia, no parecen ser tan básicos y corporizados como tradicionalmente se los ha concebido en la literatura canónica.

En cualquier caso, es probable que los sujetos apliquen tanto los esquemas miméticos como los de imagen a lo largo de toda su vida (es decir, cualquiera de los diferentes tipos de pensamiento esquemático propuestos o requeridos por la TMC). Como ejemplo de posible aplicación de EM en la adultez, la dinamicidad gestual de abrir y elevar enérgicamente los brazos por parte de quien dirige una orquesta funciona como guía rápida y eficiente para comunicar a los demás miembros de la orquesta que el pasaje adquiere potencia (*en forte*). Los signos interpretables y las cargas evaluativas de la fig. 3 parecen seguir estando presentes en ciertos movimientos que ya no son, en el caso de la dirección de orquestas, ni enteramente básicos ni automáticos, pero son fácilmente inteligibles por los intérpretes musicales, en virtud de la expresividad y de la percatación de la importancia del movimiento del cuerpo (propio y de otros) desde la infancia, ancladas en la mimesis corporal y reelaboradas de manera exponencialmente más sofisticada en otras modalidades a lo largo de la ejecución de un pasaje musical. Así, una perspectiva ontogenética de la metáfora conceptual multimodal debería dar cuenta

de cómo surge, en la temprana ontogenia, una estructura esquemática acorde con el grado de complejidad cognitiva que es atribuible en esa etapa.

Figura 3: esquema mimético: ALCANZAR



Volvamos por un momento al ejemplo de metáfora multimodal del pasaje de Palestrina en la sección 2, que contiene un dominio espacial simple disparado por la expresión lingüística “*descendit*”, tradicionalmente asociado al esquema de imagen VERTICALIDAD en la TMC. En su lugar, consideremos ahora la aplicación del esquema mimético BAJAR en la cognición, que ofrece mayor dinamicidad y corporeidad (a través de la mimesis). Como resultado, es posible que haya mayor semejanza entre este EM con el proceso estructurado dinámica y ordenadamente por las seis voces que descienden de escala: una isomorfía en lo dinámico del movimiento del cuerpo por un lado, y del sonido por el otro, lo cual es ventajoso en términos cognitivos en cuanto podría acelerar su procesamiento.

Además, con el EM nada se pierde en términos evaluativos o afectivos, ni en términos de la incongruencia típica entre los dominios que componen las metáforas.

4. Contrastación entre esquemas de imagen y esquemas miméticos

Existen dos diferencias sustantivas entre los esquemas de imagen y los esquemas miméticos, con sus correspondientes efectos en cuanto elementos constitutivos de las metáforas conceptuales, en particular, de las metáforas multimodales cuyos dominios-fuente están anclados a una experiencia sensoriomotriz, corporizada y prelingüística y que, por lo tanto, surgirían en la temprana ontogenia –aunque persistan a lo largo de la vida-. Una divergencia es el grado de abstracción (y especificidad) del pensamiento esquemático, y otra es la direccionalidad de la subjetividad en su constitución.

La primera divergencia es con respecto de la cuestión del grado de corporeidad y de abstracción de los esquemas de imagen. En Hampe (2005), se advierten y compilan divergencias definicionales e inconsistencias teóricas, que muestran que casi la totalidad de los esquemas de imagen listados en la literatura pueden considerarse de elevada abstracción y baja especificidad (p. ej. VERTICALIDAD, TODO-PARTE, PROCESO), si se sigue el criterio de Langacker para la noción de esquematicidad. Por contraste, en los esquemas miméticos, la acción y el movimiento del sujeto necesariamente quedan anclados al movimiento del cuerpo que es percibido y al movimiento de su propio cuerpo, en forma de relación causal entre el entorno socio-cultural del infante, las emociones y prácticas compartidas socialmente, y su esquema mimético (p. ej. BESAR, SALTAR). Siguiendo las caracterizaciones de Langacker y

Cienki, los esquemas miméticos son de baja abstracción y elevada especificidad. Por su parte, Zlatev considera que los esquemas miméticos se hallan en un nivel más bajo de la cognición que los esquemas de imagen, ya que estos últimos “se forman, en parte, inductivamente, i.e. extrayendo generalizaciones a partir de propiedades compartidas de esquemas miméticos particulares” (Zlatev, 2005, p. 24). Así, la noción de esquema mimético provee un anclaje constitutivamente corporizado, en virtud del rol de la mimesis en la temprana infancia. En ese sentido, el carácter corporizado se puede aplicar a los EM de manera más consistente que lo que, de hecho, se aplica para los EI. Bajo cualquier caracterización del pensamiento esquemático, sea la provista por Langacker o la más tradicional de Lakoff y Johnson, un esquema mimético es corporizado por necesidad, ya que liga las recurrencias percibidas de actos con el cuerpo (propio o de otros) y su correspondiente estructura representacional en el pensamiento.

El otro punto de contraste refiere a la subjetividad en la constitución del pensamiento esquemático desde su origen ontogenético. Se ha postulado al EI como prelingüístico para un individuo en un entorno: a todas luces, el contenido del EI es intrapersonal y se manifiesta posteriormente como interpersonal o se hace público mediante el lenguaje una vez adquirida una lengua materna. En virtud de la capacidad humana de categorizar mediante prototipos, un individuo, como miembro de una cultura y en un momento dado, posee esta estructura básica de EI. En contraste, para los esquemas miméticos, el punto de partida es primero interpersonal y luego intrapersonal: el esquema mimético se va formando desde la exterocepción y propiocepción, a medida que el sujeto va percibiendo recurrencias significativas de actos del cuerpo o de actos con el cuerpo en su entorno social. Un esquema mimético

vincula e interconecta primeramente un contenido interpersonal, para transformarse luego en un contenido intrapersonal. La inspiración, en el caso de la esquematicidad mimética, es vygotskiana (Vygotsky y Luria, 1993). En esta noción, si bien el individuo es parte de una cultura, se hace foco en la configuración subjetiva de la estructura en diferentes estadios del desarrollo, es decir, en la evolución del pensamiento mimético hasta llegar a un estadio lingüístico (ontogénesis).

A los efectos de esta indagación sobre metáforas multimodales, es interesante que Zlatev capture, mediante la noción de EM, una idea intuitiva acerca de la cognición humana: que el movimiento corporal, la percepción del cuerpo y del entorno, la gestualidad y, hasta cierto punto, el lenguaje, coocurren desde muy temprano en la ontogenia, en virtud del mero contacto social y con el mundo. Adicionalmente, esta perspectiva podría dar cuenta de cierta tendencia a una ligera divergencia experiencial resultante de los diferentes conceptos que existen en una lengua u otra, en co-ocurrencia con movimientos corporales significativos propios de cada entorno cultural.

5. Algunos aportes de la indagación sobre multimodalidad en la TMC

Las metáforas conceptuales, como se ha mencionado ya, suponen un alejamiento de las concepciones de la metáfora que dependen de manifestaciones verbales, o que están necesariamente constituidas por manifestaciones verbales. Por el contrario, toda MC depende, en última instancia, de una cognición basada en experiencias corporizadas, las cuales dan forma o hacen inteligibles conceptos de mayor abstracción, como TIEMPO o BONDAD. Sin embargo, las MC más estudiadas han sido siempre las que se manifiestan abundante-

mente en el habla cotidiana, probablemente por ser fácilmente reconocibles, y hasta fáciles de rastrear, registrar y reproducir en formatos amigables para la indagación académica tradicional. Ahora bien, en el contexto teórico de la TMC, las metáforas multimodales son importantes en cuanto revelan formas de procesamiento cognitivo que no dependen de su manifestación exclusivamente verbal, y, con ello, evitan el riesgo de circularidad explicativa que amenaza a la teoría. En efecto, al indagar sobre las MM en el marco de la TMC en la temprana ontogenia, no se asume competencia lingüística alguna, sino una diversidad de modalidades experienciales previas o diferentes de ella. Mi hipótesis es que las metáforas multimodales pueden explicarse adecuadamente mediante la noción de esquemas miméticos, que operan desde la ontogenia temprana en adelante, y que están muy ligadas a la cognición corporizada. Considerando la ontogenia y el desarrollo cognitivo que Zlatev (2013, p. 53) postula para la aparición del pensamiento esquemático mimético (alrededor de los 14 meses de edad, durante el estadio de mimesis tríadica), para que esta estructura sea cognitiva y no “ciega”, he considerado ya en Mahler (2017) que el infante debería actuar conforme a reglas de corte deflacionario, específicamente bajo una normatividad primitiva (Ginsborg, 2011), noción que describe reglas que un infante considera apropiadas al contexto *simpliciter*. Considero que esta noción es apropiada para dar cuenta de estas formas de cognición en individuos prelingüísticos, teniendo en cuenta la temprana ontogenia y sin hacer descansar tal noción de adecuación en ninguna regla general o conceptual previamente aprendida por este individuo (Mahler, 2017, pp. 376-378). Así, el infante posee un pensamiento esquemático mimético que es una disposición compleja (ni automática o ciega, ni tampoco sofisticada conceptualmente) para actuar del mismo modo que se venía haciendo en el propio entorno afectivo y social específico. Estos actos con el cuerpo emergen como significa-

tivos en dicho entorno. En ese sentido, el infante tendría “razones” para actuar con el cuerpo del modo en que lo hace, debido a que este sujeto u otros sujetos venían actuando de esa manera, dadas específicas circunstancias. Ello no compromete al infante a tener que saber que sigue reglas, ni qué reglas sigue (es decir, tener conceptos plenos)³³, lo cual, por supuesto, no sería plausible en virtud del desarrollo cognitivo propio de la temprana ontogenia. Los EM, como representaciones de actos corporales, están entonces normados en virtud de su función comunicativa en contextos determinados.

Por un lado, como se puede vislumbrar en la contrastación entre concepción de esquematicidad de Langacker (1987, 2008) y las más tradicionales en la lingüística cognitiva, no parece ser el caso de que todo pensamiento esquemático se adecue al criterio de baja abstracción y alta corporeidad requerido por la TMC, específicamente en el caso de los EI como dominio-fuente de las metáforas conceptuales. Por otro lado, si bien Kövecses (2017) detecta el caos conceptual y terminológico que rodea a la cuestión de una adecuada caracterización y denominación de los constituyentes de la MC, intenta resolverlo incluyendo varias nociones teóricas en danza en un *continuum* que analiza en términos de mayor o menor esquematicidad y especificidad en un gradiente, y, al hacerlo, no atiende en su justa consideración el problema de la abstracción para el pensamiento esquemático prelingüístico que se forja en la temprana ontogenia. Como suele ocurrir en la lingüística cognitiva, afirma que el pensamiento esquemático de imagen “surge de nuestras experiencias corporizadas más básicas” (Kövecses, 2017, p. 20); y, más aún, pareciera que para un sujeto ya adulto solamente. Así, se intuiría que la preocupación por el desarrollo ontogenético en la conformación del pensamiento esquemático, y, por lo tanto, del metafórico, en

³³ Seguir reglas como facultad del entendimiento es explícitamente una caracterización kantiana en Ginsborg (2011, p.230).

cuanto el primero constituye al segundo, es solo nominal o, cuanto mucho, lateral. Finalmente, Cienki (2013) y Forceville y Urios-Aparisi (2009) advierten acerca del riesgo de caer en circularidad, al señalar que, si “las expresiones metafóricas verbales nos brindan evidencia de las metáforas conceptuales”, ya que estas últimas “se expresan verbalmente” (Cienki, 2013, p. 422), se deja sin explicación el aspecto cognitivo de las metáforas. Estos autores, entre otros que estudian la metáfora multimodal, prestan atención teórica a la gestualidad, la música, sonidos no verbales, por ejemplo, especialmente para los dominios-fuente, en el análisis de metáforas conceptuales multimodales, procurando eludir así cualquier riesgo de circularidad explicativa.

En términos generales, la noción de dominio-fuente tendrá prioridad explicativa sobre la estructura metafórica por varias razones. Por un lado, en cuanto elementos constitutivos necesarios de la MC, no se podría concebir una metáfora conceptual sin dominio-fuente y dominio-destino. Pero, si además es cierto que el dominio-fuente precede a la comprensión que sí se termina logrando una vez proyectado sobre el dominio-destino, y si el dominio-fuente es preconceptual y esquemático, entonces es necesario estudiar el dominio-fuente como estructurador del modo en que pensamos el dominio-destino en la MC, teniendo en cuenta el enfoque corporizado, experiencial y de efectos prototípicos que propone la TMC, junto con los más recientes aportes de los estudios sobre multimodalidad y las ciencias cognitivas relevantes. Es necesario abordar el problema de emergencia en el desarrollo cognitivo infantil debido a que es posible determinar el mecanismo, en una determinada etapa de desarrollo ontogenético, en el queemergería una primera forma de pensamiento esquemático capaz de organizar y facilitar tan eficientemente la comprensión de dominios más abstractos en las MC. Pienso, además, que del tipo de estructuración de la TMC se sigue que

existe una dependencia ontogenética de todo otro tipo de conceptualizaciones con respecto al pensamiento esquemático, puesto que el pensamiento esquemático normado primitivamente, mencionado con anterioridad, daría lugar a las demás estructuras conceptuales de sofisticación creciente³⁴, dentro de las cuales se encuentran las estructuras que contienen una proyección, como es el caso de las metáforas.

6. Conclusión

Dado que la noción de dominio es teóricamente relevante para dar cuenta de la estructura de la metáfora conceptual, es importante reconocer que estos dominios pueden manifestarse monomodal o multimodalmente en ella. A partir del análisis de los rasgos de los esquemas de imagen en el pensamiento esquemático involucrado en las MC, y, en particular, del grado variable de abstracción y de corporeidad descripto en la literatura de la lingüística cognitiva,

³⁴ He considerado y desestimado la participación de los Modelos Cognitivos Idealizados (MCI) en los dominios-fuente de las MM que surgen en la temprana ontogenia. Lakoff (1987) ha sostenido que es posible describir la estructura de los conceptos mediante MCI, que son, como los EI, estructuras gestálticas. Los MCI contienen los significados de los conceptos y también se basan en la experiencia corporizada. Si no se basan de modo directo en la experiencia corporizada, lo hacen en la imaginación. Los principios que organizan la información dentro de este clúster son de mucha mayor complejidad que los del pensamiento esquemático, ya que no solamente incluyen este último, sino que además organizan el conocimiento mediante estructuras proposicionales y proyecciones metonímicas y metafóricas. En este sentido, este tipo de estructura cognitiva posibilita un entramado más sofisticado de experiencias culturales, que pueden sostener metáforas conceptuales que no dependen directamente de la experiencia corpórea: la metáfora conceptual estructural LA DISCUSIÓN ES UNA GUERRA contiene un dominio-fuente (GUERRA) que, en el mejor de los casos, no se conforma mediante una vivencia directa, pero que refiere a un conjunto de proposiciones (datos, información), sensaciones y experiencias que conocemos fácil y detalladamente de modo indirecto, incluso inferencialmente. Es fundamental destacar que este tipo de estructura cognitiva no puede emergir temprano en la ontogenia, aunque sí es posible afirmar que sus mecanismos de efectos prototípicos de percepción y semejanza tengan origen en el pensamiento esquemático.

he defendido también que los esquemas miméticos (Zlatev, 2002; 2005; 2007a; 2013) normados primitivamente (Mahler, 2017), y no los esquemas de imagen, resultan una mejor alternativa para elucidar la estructuración de las metáforas multimodales cuyo dominio-fuente requiere de un vínculo muy cercano con la diversidad de la experiencia sensoriomotriz, sin excluir su co-ocurrencia con el lenguaje. Los esquemas miméticos poseen un anclaje constitutivamente corporizado, a diferencia de los esquemas de imagen, que no pueden caracterizarse como corporizados desde todas las perspectivas de la lingüística cognitiva de manera consistente. Además, el interés por la perspectiva diacrónica u ontogenética en cuanto al rol o caracterización del pensamiento esquemático como constitutivo de las MC deriva de la conexión entre el entorno sociocultural y los factores cognitivos a lo largo del desarrollo del infante, a partir de habilidades anteriores y diferentes de las capacidades sofisticadas que entraña ser un sujeto con maduración biológica y desarrollo cognitivo plenos. Por otra parte, he defendido también la prioridad explicativa del pensamiento esquemático mimético sobre todo otro tipo de conceptualización en los dominios-fuente para dar cuenta de la estructura metafórica, y la posible dependencia ontogenética de todo otro tipo de conceptualización con respecto a estos esquemas. La hipótesis defendida acerca de la estructuración de las metáforas multimodales cuyo dominio-fuente es de tipo esquemático mimético robustece el enfoque corporizado y evita un riesgo de circularidad explicativa en la TMC, por no recurrir ni depender de manifestaciones verbales exclusivamente.

Referencias bibliográficas

- Cienki, A. (2013). "Image and mimetic schemas in cognitive linguistics and gesture studies". En M. J. Pinar Sanz (Ed.), *Multimodality and Cognitive Linguistics* (pp. 417-432). Ámsterdam: John Benjamins.
- Eameo (2018). Recuperado de <https://goo.gl/BtEFFv>.
- Forceville, C. y Urios-Aparisi, E. (eds.) (2009). *Multimodal Metaphor*. Berlín: Mouton de Gruyter.
- Ginsborg, H. (2011). "Primitive Normativity and Skepticism about Rules". *The Journal of Philosophy*, tomo CVII, n.º 5, pp. 227-254.
- Hampe, B. (Ed.). (2005). *From Perception to Meaning: Image Schemas in Cognitive Linguistics*. Berlín: Mouton de Gruyter.
- Kant, I. ([1781] 1968). *Critique of Pure Reason*. Trad. N. K. Smith. Nueva York: St. Martin's Press.
- Kövecses, Z. (2017). *Levels of Metaphor*. Berlín: Mouton de Gruyter.
- Lakoff, G. y Johnson, M. ([1980] 2009). *Metáforas de la vida cotidiana*. Madrid: Teorema.
- Lakoff, G. y Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and Its Challenge to Western Thought*. Nueva York: Basic Books.
- Lakoff, G. y Turner, M. (1989). *More Than Cool Reason: A Field Guide to Poetic Metaphor*. Chicago: University of Chicago Press.
- Lakoff, G. (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.
- Lakoff, G. (1990). "The Invariance Hypothesis". *Cognitive Linguistics*, 1(1), pp. 39-74.

- Lakoff, G. (1993). "The Contemporary Theory of Metaphor". En A. Ortony (ed.). *Metaphor and Thought* (2.º ed.) (pp. 202-251). Nueva York: Cambridge University Press.
- Langacker, R. (1987). *Foundations of Cognitive Grammar*. Vol. 1. Stanford: Stanford University Press.
- Langacker, R. (2008). *Cognitive Grammar: A Basic Introduction*. Nueva York: Oxford University Press.
- Mahler, C. (2017). "Esquemas miméticos y normatividad". En M. Bortolón, E. Pérez, M. Montes y P. García Ficarra (Comp.). *Aportes a la lingüística cognitiva* (pp. 365-380). Córdoba: Editorial de la UNC.
- Orchard Music (en nombre de coro), y 1 Music Rights Societies (productores). (2012). *Missa Papae Marcelli: Credo*. The Sixteen, Harry Christophers (intérpretes) y Giovanni Pierluigi da Palestrina (autor) [1990]. Recuperado de <https://goo.gl/PZpDtp>.
- Prinz, J. (2002). *Furnishing the Mind. Concepts and Their Perceptual Basis*. Cambridge, Mass.: MIT Press.
- Ricoeur, P. (1978). "The Metaphorical Process as Cognition, Imagination, and Feeling". *Critical Inquiry*, 5(1), pp. 143-159.
- Rosch, E. (1978). "Principles of categorization". En E. Rosch y B. Lloyd (eds.), *Cognition and Categorization* (pp. 27-48). Hillsdale, Nueva Jersey: Lawrence Erlbaum.
- Turner, M. (1990). "Aspects of the Invariance Hypothesis". *Cognitive Linguistics*, 1(2), pp. 247-256.
- Vygotsky L. S. y Luria, A. R. (1993). *Studies in the History of Behaviour. Ape, Primitive, and Child*. Nueva Jersey: Lawrence Erlbaum.

- Zlatev, J. (2002). "Mimesis: The 'missing link' between signals and symbols in phylogeny and ontology?". En *Mimesis, Sign and Language Evolution*, Anneli Pajunen (Ed.), Publications in General Linguistics 3, University of Turku, Finland.
- Zlatev, J. (2005). "What's in a schema? Bodily mimesis and the grounding of language". En B. Hampe (Ed.), *From Perception to Meaning: Image Schemas in Cognitive Linguistics* (pp. 313-342). Berlín: Mouton de Gruyter.
- Zlatev, J. (2007a). "Embodiment, language, and mimesis". En T. Ziemke, J. Zlatev y R. Frank (Eds.), *Body, Language and Mind, vol. 1: Embodiment* (pp. 297-337). Berlín: Mouton de Gruyter.
- Zlatev, J. (2013). "The mimesis hierarchy of semiotic development: Five stages of intersubjectivity in children". En *The Public Journal of Semiotics* 4(2), pp. 47-73.

Parte II. Percepción directa de otras mentes

Comportamientos expresivos naturales y contenidos mentales no proposicionales

LAURA DANÓN (CIFFYH – UNC)

1. La triangulación intermedia y el comportamiento expresivo natural

En una serie de artículos que han sido extensamente discutidos, Donald Davidson (1982, 1999, 2001a, 2001b) ha defendido que, para poder tener estados mentales intencionales con contenidos objetivos, una criatura debe, necesariamente, ser partícipe de una relación de triangulación básica que involucre tres elementos: ella misma, otra criatura similar y un objeto público situado en un espacio común. En este tipo de interacción, cada una de las criaturas debe responder tanto al objeto situado en el espacio público, como al otro participante. Pero, además, cada una ha de ser capaz de correlacionar las respuestas del otro participante con aquel objeto del mundo al cual ella también está reaccionando.

Davidson piensa que este tipo de interacción tríadica constituye una condición necesaria, pero no suficiente, para la emergencia del pensamiento. La aparición de genuinos estados mentales, dotados de contenidos objetivos, solo puede tener lugar cuando las criaturas que participan en la situación de triangulación comienzan a interactuar de un modo específico: interpretándose unas a otras

lingüísticamente. Muy sucintamente, la razón de ello es que solo cuando somos capaces de entender las emisiones lingüísticas ajenas logramos dominar los conceptos de verdad y error, y, con ello, podemos comenzar a tratar a los demás como sujetos que tienen un punto de vista, que puede ser correcto o no, acerca de un entorno común. Todo lo cual, piensa Davidson, es crucial para contar con pensamientos objetivos. El problema, a su entender, es que también se da la relación de dependencia inversa: para interpretar lingüísticamente a otros es preciso contar con pensamientos objetivos. Por esto, el dominio lingüístico y la capacidad para tener pensamientos con contenidos objetivos resultan mutuamente interdependientes: o una criatura cuenta con ambas capacidades o no tiene ninguna.

Estas ideas se complementan armónicamente con la tesis davidsoniana de que existe un abismo explicativo infranqueable entre el comportamiento no cognitivo y arracional de las criaturas carentes de lenguaje y las acciones, racionales e intencionales, que son capaces de desplegar los sujetos humanos que poseen tanto pensamientos objetivos como capacidades lingüísticas. Mientras que las criaturas carentes de lenguaje -sean estos niños prelingüísticos o animales no humanos- solo pueden participar de las formas básicas y primitivas de triangulación no lingüística, los animales humanos lingüísticamente competentes pueden involucrarse en variantes “reflexivas” o “lingüísticas” de triangulación (Bar-on y Priselac, 2011). Lo que no será posible encontrar es alguna suerte de “espacio intermedio” que conecte estos dos tipos de interacción triangular y nos ayude a explicar el pasaje de una a la otra³⁵.

³⁵ No voy a discutir aquí los argumentos que ofrece Davidson para defender la necesidad de participar de una forma de triangulación pura o básica para adquirir contenidos objetivos. Tampoco me detendré en sus breves señalamientos

Es posible describir y explicar, apelando al vocabulario de las ciencias naturales, las respuestas discriminatorias de las criaturas carentes de lenguaje. De modo similar, cabe apelar al vocabulario intencional para dar cuenta de las acciones de los animales humanos dotados de mente y de lenguaje. Pero no es posible trazar los pasos intermedios que llevan del primero al segundo, y, debido a ello, no es posible tampoco explicar cómo pudo surgir el pensamiento objetivo en un mundo natural.

En contra de tal enfoque discontinuista, Bar-on y Priselac (Bar-on y Priselac, 2011; Bar-on, 2013, 2018) defienden que es posible identificar en ciertas criaturas no lingüísticas un tipo de triangulación intermedia destinada a cubrir, siquiera parcialmente, la brecha entre las otras dos formas de triangulación postuladas por Davidson. Esta nueva variante de triangulación se caracteriza, a su entender, por un modo diferente de interacción que no se reduce a las respuestas no cognitivas que pueden dar las criaturas insertas en formas de triangulación pura, pero tampoco involucra el despliegue del tipo de capacidades cognitivas sofisticadas que se ponen en juego en los casos de triangulación lingüística.

La necesidad de postular esta suerte de triangulación intermedia se vuelve patente, según estos autores, cuando atendemos a un tipo específico de conductas que los animales humanos compartimos con diversas especies no humanas: *los comportamientos expresivos naturales*. Estos comportamientos, argumentan, presentan rasgos distintivos que no pueden ser explicados adecuadamente en los términos puramente causales que empleamos para dar

sobre por qué este tipo de triangulación básica resulta insuficiente para la emergencia de pensamientos genuinos. En Verheggen (2013), Gluer (2009), Pagin (2001) y Ludwig (2011), se pueden encontrar algunas reconstrucciones de estos argumentos, así como objeciones realizadas a ellos.

cuenta de las respuestas de las criaturas involucradas en situaciones de triangulación pura. Sin embargo, tampoco parecen involucrar las capacidades cognitivas y lingüísticas sofisticadas, asociadas al pensamiento proposicional pleno, que ponen en juego las criaturas lingüísticamente competentes en las situaciones de triangulación lingüística o reflexiva (Bar-on y Priselac, 2011). En tal sentido, pueden ser considerados como *precursores naturales* de formas más avanzadas, y distintivamente humanas, de comunicación (Bar-on, 2013).

Como ocurre en los otros dos tipos de triangulación, la variante intermedia involucra la interacción entre dos criaturas similares y un objeto del entorno común al cual ambas responden. Pero, a diferencia de lo que sucede en los casos de triangulación primitiva, los participantes de este tipo de interacción no despliegan meras respuestas discriminatorias a los objetos de su entorno, sino genuinos comportamientos expresivos. La hipótesis central de estos autores es que, en los casos de triangulación intermedia, los comportamientos expresivos naturales que realiza uno de los participantes *muestran (show)* sus estados mentales –en lugar de describirlos–, por lo cual estos se vuelven perceptualmente reconocibles para la otra criatura involucrada en la interacción, bajo el supuesto de que esta última cuenta con la dotación cognitiva adecuada.

Un punto central que estos autores recalcan es que la relación de expresión que tiene lugar en todos estos casos es diferente de la que Sellars (1969) denominaba una “expresión en el sentido semántico” o una “s-expresión”. La relación de s-expresión es un vínculo semántico, representacional y convencional, que tiene lugar entre cadenas lingüísticas, o algún otro tipo de representaciones, y aquello a lo que estas refieren. La relación de expresión que tiene lugar en los casos de comportamientos expresivos,

en cambio, es bautizada por Bar-on y colegas como una “m-expresión”. A su entender, este tipo de expresión tiene lugar cuando una criatura *muestra* o *vuelve observables* algunos de sus estados mentales, ya sea voluntaria o involuntariamente, a través de sus conductas corporales, gestos, expresiones faciales, etc., de un modo tal que, para los observadores adecuados, dichos estados se vuelvan perceptibles de modo directo (Bar-on y Priselac, 2011; Bar-on, 2013; Bar-on y Green, 2010). Esto mismo diferencia los comportamientos expresivos naturales de lo que Grice denominaba “signos naturales”: esto es, indicadores confiables de aquellos estados mentales que regularmente son su causa. Un signo natural, como las manchas rojas en la piel de quien sufre sarampión, provee evidencia más o menos confiable de la presencia de algo. Un comportamiento expresivo –como un aullido, una sonrisa, un grito de alarma, etc. – puede ser caracterizado, en cambio, como

una forma de comportamiento público, social, intersubjetivo, dirigido hacia el mundo, que ha sido diseñado por la naturaleza (por oposición a haber sido diseñado por la cultura o por las intenciones de un individuo), para mostrar o exhibir la presencia y características de los estados mentales de quien así se expresa ante ciertos observadores adecuadamente dotados (Bar-on, 2018, p. 193).

Un rasgo interesante de los comportamientos expresivos naturales –como el mostrar los dientes, gruñir, sonrojarse, cerrar los puños, etc.– es que no siempre pueden ser equiparados a las acciones que se llevan a cabo de modo voluntario e intencional (Bar-on y Green, 2010). Antes bien, a menudo cabe describirlos como comportamientos que realizamos de manera automática y no plenamente consciente. En algunos casos puede tratarse, incluso, de comportamientos que tienen lugar *en contra* de la voluntad de quien los realiza. Luego, un comportamiento

expresivo puede hacer manifiesto cuál es el estado mental en que se encuentra el animal, sin que este se haya formado una intención de revelar tal cosa a los demás.

Por otra parte, los comportamientos expresivos naturales no se limitan a exhibir o manifestar los estados mentales de quien los realiza. También nos proporcionan información sobre las acciones que este se apronta a efectuar en un futuro inmediato, e instan al observador a actuar de un modo específico. Para entender mejor esto, podemos imaginar una situación de triangulación intermedia en la cual dos criaturas, C1 y C2, observan un mismo objeto en el entorno común –en este caso, un predador–, así como también sus mutuas respuestas a él. Supongamos que, en este contexto, C1 da un grito de alarma ante la presencia del predador. Bajo el marco interpretativo que se nos propone, dicho grito operaría como un comportamiento expresivo que manifiesta un estado mental de C1, su temor ante el predador, e indicaría a la audiencia adecuada –en este caso, a C2– que C1 está por huir apresuradamente, instándola a comportarse de un modo similar.

Un punto adicional que señalar es que, en la medida en que el comportamiento expresivo de C1 proporciona información sobre cómo va a responder a O en un futuro inmediato, C2 está en posición de responder de un modo que tiene en cuenta no solo la presencia de O, sino también la respuesta que anticipa en C1. En este contexto, puede ocurrir que C2 termine dando una respuesta a O que no se ajuste a la que dará C1. Así, por ejemplo, puede que C2 elija no huir del predador –porque estima que aún no está suficientemente cerca–, sino quedarse un rato más para obtener el alimento de C1 una vez que este se haya alejado. De acuerdo con Bar-on y Priselac, en un caso como este, la respuesta discordante de C2 manifiesta su desacuerdo, conductualmente encarnado, en relación con cómo se ha

de responder a O. Incluso se podría pensar, añaden estos filósofos, que C2 está tratando las respuestas de C1 como algo *distinto* y *separable* de la presencia misma de O. Pero, si C2 es, efectivamente, capaz de trazar una separación de tal tipo, entonces puede, en cierto sentido, tratar a C1 como teniendo su propia comprensión de la situación en la que ambos se encuentran cuando enfrentan al objeto O, y posee alguna sensibilidad, siquiera en la práctica, a la posibilidad de que C1 se equivoque con respecto al tipo de respuesta que ha de darse ante él (Bar-on y Priselac, 2011; Bar-on, 2013, 2018). Esta es, sin duda, una de las razones principales por las que cabe considerar que este tipo de triangulación ocupa un genuino terreno intermedio entre los casos de triangulación pura, en los que, a lo sumo, C2 puede responder ante O con un comportamiento que no se corresponde con el de C1, y los casos de triangulación reflexiva, en los que C2 explícitamente juzga el comportamiento y los estados mentales de C1 como correctos o incorrectos (Bar-on, 2013, 2018).

Esta no es, sin embargo, la única diferencia relevante entre los casos de triangulación intermedia y los de triangulación lingüística. Bar-on y colegas también han defendido consistentemente una tesis controvertida con respecto al contenido de los estados mentales que una criatura manifiesta mediante sus comportamientos expresivos naturales. Su idea es que tales estados no han de ser caracterizados, del modo clásico, como actitudes proposicionales, sino como estados de índole *no proposicional* (Bar-on, 2010; Bar-on y Priselac, 2011; Bar-on y Green, 2010; Bar-on, 2013, 2018). No es claro, sin embargo, cuál es el modo específico en que debemos entender tales afirmaciones, ya que no encontramos en sus textos ni una teoría detallada propia sobre la naturaleza de los contenidos no proposicionales, ni una clara adhesión a alguna de las

teorías alternativas sobre el tema que cabe encontrar en la literatura. El próximo apartado, por lo tanto, estará dedicado intentar esclarecer el alcance que tiene, para estos autores, tal noción.

2. El contenido de los comportamientos expresivos naturales: dos alternativas

En lo que sigue, querría focalizarme en el problema principal de este trabajo: ¿qué tipo de contenidos tienen los estados mentales que manifiestan, mediante sus comportamientos expresivos, las criaturas que participan en situaciones de triangulación intermedia? En distintas ocasiones, Bar-on y colegas señalan que tales estados mentales se diferencian de las actitudes proposicionales clásicas por poseer contenidos de naturaleza no proposicional (Bar-on y Priselac, 2011; Bar-on, 2010, 2018, 2013; Bar-on y Green, 2010). Sin embargo, hay modos muy diferentes en los que los filósofos han entendido la dicotomía *proposicional/no proposicional*, y, aunque los autores ofrecen algunos señalamientos generales acerca de los rasgos distintivos que tendrían los contenidos mentales que los ocupan, resulta difícil precisar cuál es exactamente la noción de contenido no proposicional a la cual suscriben. Dadas estas dificultades, en lo que sigue querría detenerme a examinar sus afirmaciones sobre el tema, con el objetivo de deslindar de qué modo hemos de entender los contenidos no proposicionales de los estados mentales primitivos que se expresan (y perciben) en las situaciones de triangulación intermedia.

Según vimos anteriormente, una de las tesis centrales que defienden Bar-on y colegas es que los comportamientos expresivos de una criatura C manifiestan sus actitudes

psicológicas -por ejemplo, su temor, su alegría, etc.- de un modo tal que un observador con la adecuada dotación cognitiva debería poder reconocerlas de un modo perceptual y no inferencial. Usualmente, añaden a esto que muchos de estos comportamientos expresivos naturales -los gestos, la orientación general del cuerpo, la dirección de la mirada, etc.- manifiestan, además, cuáles son los objetos a los cuales se dirigen tales actitudes psicológicas (Bar-on y Green, 2010; Bar-on y Priselac, 2011; Bar-on, 2018, 2013). Si, por tomar un caso posible, C1 diera un grito de alarma, comenzara a temblar, fijara la mirada en un predador cercano, etc., cabría concluir que su comportamiento expresa su temor *a ese predador*.

En lo siguiente, querría sugerir que se puede considerar que estas afirmaciones van aproximando a Bar-on y colegas a la tesis de que los estados mentales que una criatura expresa mediante ciertos comportamientos expresivos naturales son estados dotados de *contenidos de objeto*, siguiendo, siquiera *grosso modo*, los desarrollos teóricos de filósofos como Crane (2013), Montague (2007) y Grzankowski (2013, 2016, 2017). Para quienes defienden la noción de contenidos de objeto, hay ciertos estados mentales -como *amar a Dios*, *temer a la oscuridad* o *desear chocolate*- que involucran una relación entre un sujeto y un objeto particular. Grzankowski (2016) añade que, en tales casos, lo que el sujeto tiene en mente es algo similar a una palabra mental -en lugar de una oración mental-, esto es, una representación carente de condiciones de verdad. Su contenido, por lo tanto, es no proposicional y de carácter objetual.

Al menos en primera instancia, hay dos maneras posibles de entender los contenidos de objeto: como contenidos que versan acerca de algún objeto *simpliciter* -sin involucrar un modo específico de caracterizarlo- o como

contenidos que representan un objeto, pero de un modo aspectual. En el primer caso, lo que el comportamiento expresivo de C revelaría a una audiencia en posesión de una dotación cognitiva adecuada sería, a lo sumo, la actitud psicológica de C, más una especie de señalamiento *de re* al objeto de esta. Si así fueran las cosas, el comportamiento de C podría mostrar a su audiencia, por ejemplo, que C teme a la araña, pero no le daría ninguna información adicional con respecto a cómo C se representa la araña, o de qué modo la caracteriza y categoriza. Dicho de otro modo, el comportamiento de C no brindaría información sobre un modo, que excluya a otros, de representar a ese objeto. Esta posición casa bien con la idea de que lo que una criatura tiene en mente, cuando alberga estos contenidos no proposicionales, es el análogo mental de un nombre, al menos bajo una concepción estrictamente referencialista de los nombres. En contraste, en el segundo caso, el comportamiento de C expresaría no solo cuál es el objeto de sus estados mentales, sino también cómo C categoriza a dicho objeto, bajo qué aspectos lo piensa, etc. Así, por ejemplo, al observar el comportamiento expresivo de C, la otra criatura que participa de la triangulación intermedia puede llegar a saber no solo que C tiene miedo de la araña, sino también que se la representa como una araña, como un objeto que se mueve, como venenosa, etc.

Aunque muchos filósofos piensan que todo contenido debe ser aspectual, la idea de que los comportamientos expresivos solo revelan las actitudes psicológicas de una criatura más sus objetos *simpliciter* tiene su atractivo. Al menos en primera instancia, parece relativamente sencillo explicar cómo, en una situación de triangulación intermedia, C2 puede llegar a adquirir información sobre cuál es el objeto *simpliciter* de la actitud intencional de C1: basta para ello con que, de algún modo, sus gestos de temor,

alegría, etc., logren direccionar la atención de C2 hacia cierto objeto del entorno común. Parece mucho más difícil, en cambio, explicar cómo los comportamientos expresivos de C1 pueden darnos información adicional acerca del aspecto específico bajo el cual se representa al objeto en cuestión. ¿Cómo podrían los gestos, postura corporal, etc., de C1 proporcionar información acerca de los modos específicos en los que se representa a O?³⁶ Por supuesto, C1 podría *contarle* a C2 cómo se representa a O; pero, si hiciera esto, y C2 pudiera entenderla, entonces ambas estarían participando, de hecho, de una forma de triangulación lingüística.

Aunque todas estas objeciones me resultan atendibles, también creo que hay modos de subsanarlas, siquiera parcialmente. Una alternativa es pensar que, una vez que C2 logra dirimir, a partir del comportamiento expresivo de C1, cuál es el objeto O al cual responde, podría sencillamente *proyectar a C1*, por *default*, el aspecto específico bajo el cual ella misma (C2) representa a O. Luego, si C2 ve a O como un predador peligroso, puede asumir que C1 también lo representa del mismo modo. Esta hipótesis inicial puede verse corregida, eventualmente, por el comportamiento posterior de C1 en relación con O. Si C1 se acercara confiada a O, quizás esto bastaría para que C2 revisara su posición original.

Una segunda respuesta posible es la siguiente: aun si C2 no proyectase sus propios recortes aspectuales de los objetos sobre C1, en sentido estricto, el comportamiento expresivo de C1 hacia O podría bastar para que C2 detectase, siquiera de modo grueso, bajo qué aspectos caracteriza

³⁶ Davidson (1975, 1982, 1999) es uno de los filósofos que de modo más insistente e incisivo ha sostenido que cualquier comportamiento no lingüístico es, en último término, demasiado toscos para permitirnos atribuir contenidos mentales aspectuales. En Danón (2016), intento dar una respuesta a estos argumentos.

a O. C1 podría realizar gestos que revelasen, por ejemplo, confianza, interés, temor, aprehensión, etc. Todos ellos permitirían a C2 comenzar a delimitar bajo qué aspectos es representado O por parte de C1. Estas primeras hipótesis previas podrían verse complementadas y enriquecidas por el modo en que C1 se sigue comportando en relación con O posteriormente (¿se acerca o se aleja de O?, ¿intenta comerlo, atacarlo, huir de él?, etc.), cómo reacciona ante objetos de la misma índole (¿se eriza la piel de C1 ante O del mismo modo en que se eriza ante otros objetos peligrosos?), cómo reacciona a O cuando algunas de las propiedades de este varían, etc. Puede que, por estas vías, no logremos más que un acceso toscos y parcial al conocimiento de los aspectos bajo los cuales C1 representa al objeto al que está respondiendo. Pero esto no es lo mismo que afirmar que el comportamiento expresivo de una criatura solo nos brinda conocimiento acerca de a qué objeto responde, sin revelarnos *nada* sobre los modos en los que este es representado.

Por otra parte, con independencia de si contamos con los recursos adecuados para resolver estas dificultades, hay razones textuales para pensar que, para Bar-on y colegas, los contenidos no proposicionales en discusión sí son de carácter aspectual. Para citar un ejemplo especialmente claro, Bar-on (2018, p. 191) afirma que: “*las performances* expresivas están vinculadas a objetos y rasgos en el entorno del animal tal como estos son percibidos por dicho animal (y, en este sentido, están ‘filtradas psicológicamente’). Esta cita sugiere que, mientras que un cambio físico bruto en la criatura –como el enrojecimiento de la piel después de ser golpeada por una rama– podría ser considerado como una mera reacción ante un objeto *simpliciter*, parece más apropiado caracterizar los actos expresivos naturales como respuestas a objetos que están mediadas

por el modo en que el animal las aprehende cognitivamente. A su vez, cabe pensar que dichas “aprehensiones cognitivas” presuponen que el animal pone en juego representaciones aspectuales de los objetos en cuestión. Luego, parece que, de optar por una concepción del contenido no proposicional como un contenido de objeto, Bar-on y colegas deberían inclinarse por sostener que estos contenidos son aspectuales.

Sin embargo, como señalé anteriormente, no es enteramente claro que la noción de contenidos objetuales capture adecuada y exhaustivamente la posición de estos autores, por dos razones. En primer lugar, aunque usualmente Bar-on y colegas se limitan a afirmar que los comportamientos expresivos de una criatura nos permiten percibir cuál es el objeto de sus estados mentales, en ocasiones añaden que también pueden permitirnos conocer los *estados de cosas* a los que se refieren estos últimos (Bar-on y Green, 2010; Bar-on, 2015). Ahora bien, tradicionalmente se ha pensado que son los contenidos proposicionales, y no los objetuales, los que versan sobre estados de cosas.

En segundo lugar, cuando discuten acerca de los comportamientos expresivos en cuanto vehículos de los estados psicológicos, Bar-on y colegas con frecuencia invocan algunas ideas de Wilfrid Sellars que no resultan fáciles de reconciliar con la tesis de que los contenidos mentales que manifiestan los comportamientos expresivos son de naturaleza objetual. Según estos autores, los comportamientos expresivos constituyen vehículos carentes de una estructura articulada que permita expresar significados compasionalmente estructurados (Bar-on, 2013, 2018; Bar-on y Priselac, 2011; Bar-on y Green, 2010). A diferencia de lo que ocurre en el caso de los enunciados lingüísticos –vehículos paradigmáticos de los contenidos proposicionales–, no se puede identificar usualmente, en los patro-

nes de comportamiento expresivo, partes recombinables que operen como símbolos aislables, con un significado estable y una función sintáctica específica, y que puedan ser reutilizadas en diversas combinaciones. Sin embargo, siguiendo a Sellars, estos filósofos admiten que, aunque carezcan de estas partes, los comportamientos expresivos naturales sí poseen múltiples aspectos o dimensiones. De hecho, este último rasgo es el que les posibilita expresar estados psicológicos complejos. Más específicamente, en la versión de Sellars, esta complejidad permite que ciertos estados carentes de partes recombinables desempeñen, pese a ello, tanto la función de referir a una entidad, como la de caracterizarla como poseedora de ciertos atributos. Así pues, mientras que un enunciado representa que “*A es azul*” concatenando secuencialmente los símbolos “A”, “es” y “azul”, podemos representar este mismo hecho empleando un vehículo diferente, carente de partes, como el siguiente: un único símbolo – la letra A- que refiere al particular A, escrita en bastardillas -A- con la finalidad de representar que el particular en cuestión es de color azul (Sellars, 1981).

Curiosamente, Sellars (1981) defiende estas ideas en el marco de la introducción de un tipo de sistemas representacionales primitivos, atribuibles a los animales no humanos (De Vries, 1995). Los animales humanos, nos dice Sellars, poseerían estados representacionales dotados de una estructura sujeto-predicado, compuesta por dos tipos de símbolos separados, cada uno de los cuales cumpliría una función específica: referencial o de caracterización. Un rasgo nodal de este tipo de representaciones es que forman parte de sistemas representacionales mayores, que contienen ítems representacionales que funcionan como las conectivas lógicas y los cuantificadores (cfr. Sellars, 1981, §79). Esto posibilita a los animales huma-

nos lingüísticamente competentes formar pensamientos que involucren cuantificadores, condicionales, negaciones, disyunciones, etc. A su vez, esto les permite formular distintos tipos de principios y generalizaciones -tales como las leyes universales, los principios condicionales, etc.-, hacer transicionales inferenciales en las que figuren tales generalizaciones, y actuar a la luz de ellas. Sellars considera que el rasgo esencial de estos estados representacionales es, por lo tanto, que poseen una “forma lógica”.

Sin embargo, como anticipé arriba, este filósofo también parece pensar que existe un tipo diferente de estados representacionales primitivos, que cabe atribuir a otros animales. Él reserva el término de “proposicionales” para referirse a estos estados, en la medida en que, sostiene, comparten un rasgo central con los estados dotados de forma lógica: ellos también representan determinados objetos como poseedores de ciertas características. Con algo más de detalle, al igual que los estados representacionales con forma lógica, estos estados representacionales primitivos cumplen dos funciones semánticas diferentes: a) refieren a algo y b) caracterizan aquello a lo que refieren de cierto modo (Sellars, 1981, §72). A esto se suma que, según Sellars, para poder albergar estados representacionales primitivos que representen algo como siendo de cierto modo, es preciso que una criatura pueda realizar ciertas inferencias que conecten tal estado con otros. Luego, los estados representacionales animales estarían dotados de funciones proposicionales primitivas solo en la medida en que “funcionan como (análogos de las) premisas en los (análogos de) razonamientos prácticos que dan forma a las interacciones de los animales con sus entornos” (Rosenberg, 2007, p. 110)³⁷.

³⁷ Ver también Sellars (1981, §58).

Ahora bien, según Sellars, las funciones semánticas de referencia y caracterización no requieren, para su implementación, de una estructura oracional. Tanto el referir como el caracterizar son roles funcionales que pueden ser desempeñados por diferentes tipos de vehículos complejos (cfr. Sellars, 1981, §61). Su adhesión a esta tesis le permite explicar por qué mientras que las representaciones proposicionales primitivas carecen de sujetos y predicados (pues no es posible identificar en ellas símbolos específicos y separables que representen a uno y otro), aún pueden referir a particulares y atribuirles propiedades (cfr. Sellars, 1981, §72). Ahora bien, ¿qué es lo que permite a estos estados cumplir tales funciones aun cuando carecen de la estructura característica de los enunciados lingüísticos? Es precisamente para responder a este interrogante para lo que Sellars presenta la idea, también recuperada, como vimos, por Bar-on y colegas, de que un vehículo puede ser complejo aun si carece de partes componentes y de una estructura lógica que las articule de ciertos modos específicos. La razón de ello es que dicho vehículo podría tener múltiples aspectos, en lugar de múltiples partes, y esto le permitiría satisfacer distintas funciones semánticas (Bar-on y Green, 2010). Ahora bien, es claro que esta tesis, que luego adoptarán Bar-on y colegas, es presentada originalmente por Sellars para apuntalar una noción de estados mentales primitivos muy diferente de la de estados mentales con contenidos de objeto. Según vimos, los estados primitivos sellarsianos representarían estados de cosas, y no meros objetos. Luego, cabe preguntarse en qué medida están comprometidos Bar-on y colegas con el marco general de ideas de Sellars sobre los tipos de estados representacionales atribuibles a distintas criaturas. ¿Es acaso la noción sellarsiana de estados representacionales primitivos, en lugar de la noción de contenido de objeto, la que

subyace a sus ideas sobre el tipo de estados mentales no proposicionales que se manifestarían a través de los comportamientos expresivos?

Siguiendo el recorrido realizado hasta aquí nos encontramos con que, al tratar de esclarecer cuál es la noción de contenido no proposicional que cabe detectar en los escritos de Bar-on y colegas, hemos arribado a una posición inestable. Por una parte, algunas de sus afirmaciones sugieren que, para estos filósofos, los estados mentales que muestran los comportamientos expresivos naturales de una criatura tendrían contenidos de objeto. Sin embargo, esto entra en tensión con otras de sus afirmaciones, así como con su simpatía por ciertas tesis centrales de Sellars, que podrían llevarnos a adscribirles un tipo de contenido diferente. Ante esta situación, quizás convenga abandonar la discusión exegética sobre qué noción de contenido mental no proposicional se desprende efectivamente de sus textos, y pasar, en cambio, a preguntarnos qué noción sería más adecuado adoptar a fin de explicar el tipo de comunicación expresiva que tiene lugar en situaciones de triangulación intermedia y el tipo de capacidades conductuales y cognitivas que despliegan las criaturas carentes de lenguaje que participan de ella. Aunque este es un debate sumamente complejo, que excede los límites de este trabajo, en el siguiente apartado querría dejar esbozadas algunas consideraciones iniciales al respecto.

3. Contenidos no proposicionales: las tres vías

Según vimos en los apartados anteriores, Bar-on y colegas subrayan enfáticamente que los estados mentales que una criatura manifiesta en situaciones de triangulación intermedia, cuando responde de modo expresivo ante ciertos

objetos del entorno común, son de naturaleza no proposicional. Sin embargo, el examen del apartado anterior también nos ha mostrado que no es sencillo deslindar, a partir de sus textos, cuál es exactamente la noción de contenido no proposicional que adoptan estos autores. Más aún, según vimos, hay elementos en sus escritos que sugieren que se inclinan por una noción de contenidos mentales objetuales, mientras que, por otras razones, cabría pensar que tienen algún compromiso teórico con un tipo de contenidos primitivos *a la Sellars*. Dejando de lado los intereses exegéticos, en lo que sigue querría abordar los siguientes interrogantes: ¿cuáles son las alternativas teóricas que tendrían Bar-on y colegas para resolver esta disyuntiva? Y, de modo más relevante, ¿cuál de ellas nos brinda el mejor modo de entender los contenidos no proposicionales en el contexto de sus intereses teóricos?

Querría sugerir tres vías alternativas que los autores podrían optar por transitar.

a) La vía sellarsiana

Una primera alternativa para Bar-on y colegas consistiría en adherir plenamente a las ideas de Sellars y sostener que los gestos naturales expresivos expresan estados representacionales primitivos, que representan las cosas como siendo de cierta manera y cumplen, simultáneamente, tanto con una función referencial, como con una función de caracterización. Si este fuera el caso, el gesto expresivo de un mono que da un grito de alarma ante la presencia de un tigre, por ejemplo, expresaría un contenido como “*Hay un tigre que se aproxima*”³⁸.

³⁸ Si bien no puedo evitar emplear aquí un enunciado lingüístico para expresar el contenido que podría atribuirse a esta criatura, esto no implica que esté sosteniendo que, en el caso del animal, sus contenidos deban ser proposicionales en

Seguir esta vía suscita, sin embargo, el siguiente interrogante: ¿no sería esto equivalente a abandonar por completo la idea de que el contenido de los comportamientos expresivos naturales es no proposicional? A mi entender, la respuesta que demos a esta pregunta dependerá, en gran medida, de cómo entendamos la noción de contenido proposicional. Claramente, este último término es empleado con múltiples acepciones por parte de los filósofos. De acuerdo con la variante más deflacionada y minimalista, para que un contenido sea considerado como proposicional basta con que represente las cosas como siendo de cierta manera y, como consecuencia de ello, pueda ser evaluado como verdadero o falso, o, de modo más laxo, como correcto o incorrecto (Searle, 1982; Grzankowski, 2013; Schroeder, 2006; Crane, 2013). Al menos para los propósitos presentes, este modo de entender los contenidos proposicionales parece muy similar al del propio Sellars.

En el otro extremo del espectro, sin embargo, hay un modo pleno de entender los contenidos proposicionales, de acuerdo con el cual estos no solo representan algo como siendo de cierto modo, sino que deben contar, además, con una genuina estructura oracional. Con frecuencia esto se asocia, a su vez, con varios otros requisitos demandantes que los contenidos proposicionales deben satisfacer. En primer lugar, así como los enunciados se hallan compuestos por palabras, se piensa que los contenidos proposicionales plenos están compuestos por conceptos, o por habilidades conceptuales (Glock, *en prensa*; Burge, 2010a; Camp, 2004). En segundo lugar, debe ser posible recombi-

el mismo sentido que los nuestros, deban estar compuestos por conceptos como los nuestros, ni tampoco, más específicamente, que cuenten con conceptos análogos a los que nosotros expresamos con palabras como “tigre” o “aproximarse”. La idea, mucho más básica, es que estos animales son capaces de referir a ciertas entidades (en este caso, los tigres) y de caracterizarlas como siendo de cierto modo.

nar estos conceptos entre sí de todos los modos que resulten significativos, pudiendo formar, de esta manera, una miríada de contenidos proposicionales diferentes³⁹.

A menudo se afirma, adicionalmente, que los contenidos proposicionales deben poder incluir cuantificadores universales o existenciales. Luego, una criatura que cuenta con contenidos proposicionales ha de ser capaz de pensar contenidos como “*Existe una x tal que x es P*” y “*Para toda x, si x es p entonces x es Q*”. Suele sumarse a ello la tesis de que los contenidos proposicionales deben poder ser combinados entre sí formando unidades mayores, mediante el empleo de todo tipo de conectivas lógicas (como la conjunción, la disyunción, etc.) (Burge, 2010a, 2010b; Bermúdez, 2003; Crane, 2009). Finalmente, a veces se piensa que quien cuenta con contenidos proposicionales ha de ser capaz de separar los conceptos que constituyen tales contenidos y aprehender o captar su significado con independencia de su aplicación a cualquier situación particular (Camp, 2015; Cussins, 1992).

Ahora bien, si tomamos en consideración que existen modos tan diversos de entender los contenidos proposicionales, parece claro que buena parte de la discusión con respecto a la naturaleza de los contenidos mentales que se manifiestan mediante comportamientos expresivos naturales va a depender de cuán exigente sea la noción de contenido proposicional a la cual uno suscriba. Así, si uno adhiere a una noción demandante, tenderá a considerar que lo que Sellars denomina “contenidos proposicionales” en realidad no merecen tal nombre, en la medida en que no satisfacen todos los requisitos para ser calificados como tales. En otras palabras: si se adhiere a una concepción plena o demandante de los contenidos proposicionales,

³⁹ En otras palabras, las criaturas que poseen contenidos proposicionales han de satisfacer lo que Evans (1982) llamó el “requisito de generalidad” (Camp, 2004).

cabe reservar el término proposicional para lo que Sellars llamaba “contenidos lógicos” y caracterizar como “no proposicionales” al tipo de contenidos primitivos atribuibles a los animales que él denominaba “proposicionales”. De adoptar una estrategia de tal tipo, Bar-on y colegas podrían revisar la terminología de Sellars, sin abandonar sustancialmente el núcleo de su propuesta.

Esta alternativa, sin embargo, no carece de costos. Después de todo, estos filósofos parecen estar interesados en capturar el tipo de estados mentales que podrían poseer y expresar, en contextos de triangulación intermedia, criaturas carentes de lenguaje y relativamente primitivas desde un punto de vista cognitivo. Más aún, sus textos abundan en ejemplos de criaturas de esta índole, que meramente responden con cierta actitud psicológica ante la presencia de un objeto. Ahora bien, consideremos la siguiente conjetaura. Supongamos que, en uno de estos casos primitivos, cada vez que una criatura reconoce y categoriza a un mismo tipo de predador como tal, siente temor. Más aún, imaginemos que esto ocurre de modo sistemático y general, aun cuando la criatura no predique del predador ninguna propiedad adicional que luego pueda generalizar a otros casos, ni se represente un estado de cosas complejo, entre otros posibles, del cual el predador forme parte. Si esta conjetaura fuera correcta, cabría esperar que esta criatura desplegará, básicamente, el mismo patrón de respuestas afectivas, conductuales y cognitivas ante la presencia del objeto que ella categoriza como “predador”, con fuerte prescindencia de cómo variaran otras propiedades de este (sin importar, por ejemplo, si dicho predador se acercara por tierra o aire, si estuviera quieto o corriera, si fuera de mayor o menor tamaño, etc.). Aunque el punto merece mayor desarrollo, pienso que, en una situación muy básica de este tipo, lo más adecuado podría ser concluir,

sencillamente, que nuestra criatura alberga un estado de temor *al predador*, compuesto solo por una actitud psicológica o emocional y una representación de dicho particular. Pero esta posibilidad explicativa se pierde, pienso, si adherimos estrictamente al marco sellarsiano y abandonamos la noción de contenidos no proposicionales de objeto.

b) La vía de los contenidos objetuales

Una opción diferente consistiría en adherir estrictamente a la idea de que los contenidos mentales que una criatura manifiesta en situaciones de triangulación son estrictamente objetuales, en lugar de versar, *a la Sellars*, sobre estados de cosas y sus condiciones de satisfacción. Adoptar tal vía no necesariamente tendría que conllevar un abandono de la totalidad de las ideas de Sellars. De hecho, una estrategia que explorar sería la de emplear sus ideas sobre los vehículos complejos, pero carentes de una estructura oracional, para un nuevo fin teórico: explicar de qué modo podrían los comportamientos expresivos naturales vehiculizar contenidos de objeto dotados de formas aspectuales.

Una alternativa de este tipo podría dar lugar a un modo novedoso de entender los tipos de estados mentales que las criaturas carentes de lenguaje pueden formar y expresar en situaciones de triangulación intermedia. Sin embargo, adherir a la idea de que estos estados mentales tienen exclusivamente contenidos de objeto es una propuesta que presenta sus propios límites explicativos. Mi idea en relación con este punto es la siguiente: en algunos casos, el mejor modo de explicarlas conductas de ciertas criaturas no lingüísticas es sosteniendo que los estados mentales primitivos que estas poseen tienen contenidos que versan sobre estados de cosas y no meramente sobre objetos.

Permitásemelaborar este punto, volviendo a nuestro ejemplo de la criatura que enfrenta a un predador. Imaginemos, en este caso, que C teme al predador aéreo P, pero solo cuando lo ve aproximarse volando hacia él. Cuando, en cambio, C se encuentra ante un ejemplar de P con un ala rota, se muestra cauto al comienzo y examina con cuidado la situación. Sin embargo, una vez que comprende que P es incapaz de levantar vuelo, deja de mostrar temor. Es interesante advertir que C se comporta con cierta cautela en los dos casos, lo cual sugiere que en ambos categoriza a P como un predador. Sin embargo, su conducta sugiere que también es sensible a si P posee o no ciertas propiedades: a saber, si este puede volar o es incapaz de hacerlo por tener el ala rota. En situaciones como estas, pienso, no basta con atribuir a la criatura contenidos de objeto, pues esto no nos permitiría explicar el modo en que sus respuestas a un mismo tipo de objeto, que siempre es categorizado adecuadamente del mismo modo, varían a medida que cambian algunas de sus propiedades. La solución más simple, pienso, consiste en pensar que, en estos casos, nuestra criatura no solo se representa un predador P, sino que representa distintos estados de cosas. O, en otras palabras, se representa a P como poseedor de ciertas propiedades (propiedades adicionales, obviamente, a aquellas en virtud de las cuales fuera categorizado inicialmente como P). Adherir a la tesis de que los únicos contenidos que pueden manifestar los comportamientos expresivos son objetuales nos impide *a priori* atribuir este tipo de estados mentales, y nos deja sin la posibilidad de explicar algunos de los patrones de respuestas más sutiles y versátiles que algunas criaturas prelingüísticas puedan presentar ante su entorno.

c) La defensa de un enfoque híbrido

A la luz de las consideraciones previas sobre las limitaciones explicativas con las que tropezaríamos al optar exclusivamente por las vías a) o b), convendría sopesar una tercera alternativa: adherir a una *posición híbrida*, de acuerdo con la cual los comportamientos expresivos naturales pueden expresar, en distintas ocasiones, estados mentales con contenidos de tipos diferentes. En algunos casos, puede bastar con postular estados mentales con meros contenidos de objeto. En otros, en cambio, parece preciso pensar que tales contenidos versan sobre estados de cosas, o representan las cosas como siendo de cierta manera, y tienen condiciones de verdad o de corrección. Ahora bien, una vez que hemos aceptado la necesidad de incorporar contenidos que representan las cosas como siendo de cierta manera, la alternativa más deflacionada parece ser sostener que estos son contenidos representacionales primitivos (como los que Sellars estaba dispuesto a atribuir a los animales no humanos), que no calificarían como genuinamente proposicionales bajo una acepción demandante del término. Esta posición híbrida parece mejor posicionada para dar cuenta del tipo de dificultades y limitaciones explicativas con que se encontrarían las primeras dos vías alternativas.

Sin embargo, aún queda mucho por hacer si uno está interesado en desarrollar un enfoque de tal tipo. A modo de cierre, querría señalar dos puntos problemáticos que merecerían especial atención.

En primer lugar, parece que si uno quiere desarrollar un enfoque híbrido que integre dos tipos de contenidos primitivos -objetuales y de estados de cosas-, es preciso admitir que con ello estamos reconociendo la existencia de dos modos diferentes en los que se puede representar a algo como poseedor de ciertas características (Crane,

2013). Uno de ellos consiste en representar que algo es el caso donde tal representación puede resultar correcta o incorrecta. El otro consiste en representar algo (un objeto) bajo cierto aspecto. Pero, si esto es así, podemos preguntarnos: ¿qué es lo que distingue a estos dos tipos de contenidos?, ¿cuál es exactamente la diferencia entre la criatura que piensa un contenido objetual como “*Esa manzana!*” y la que piensa, en cambio, el contenido “*Eso es una manzana!*”?

En segundo lugar, uno puede preguntarse cuáles son exactamente los roles que cada uno de estos contenidos desempeñan en nuestra economía psíquica, y cómo se vinculan entre sí. Como distintos filósofos han subrayado, usualmente atribuimos a los estados mentales intencionales un papel clave en la explicación y predicción de los comportamientos del agente. Más aún, la ortodoxia asume que ningún estado mental puede causar comportamientos por sí mismo y de modo aislado. Antes bien, distintos tipos de estados mentales, con diferentes contenidos, han de combinarse entre sí, como parte de procesos de razonamiento práctico, a fin de llevar a las criaturas que los poseen a actuar de un modo u otro. Procesos inferenciales similares nos permiten explicar, análogamente, cómo una criatura adquiere, de modo racional, nuevos estados mentales –paradigmáticamente creencias– a partir de otros estados mentales previos.

Ahora bien, dado que contamos con los recursos lógicos para explicar, de manera estándar, las relaciones inferenciales entre proposiciones, parece que podemos luego aplicar estas mismas herramientas para dar cuenta de los distintos procesos inferenciales de una criatura, en la medida en que los contenidos de sus estados mentales cuenten con una estructura proposicional (Mandelbaum, 2012; Bermúdez, 2003).

Alguien podría cuestionar, no obstante, que lo mismo valga para el caso de los contenidos primitivos que Sellars llama “proposicionales”, pues estos carecen de forma lógica y no pueden incorporar ni cuantificadores ni conectivas lógicas. Sin embargo, en la medida en que estos contenidos representen las cosas como siendo de cierta manera, aún contamos con algunas vías teóricas para explicar cómo una criatura podría razonar apelando a ellos. Uno podría, por ejemplo, inclinarse a pensar que las criaturas que poseen contenidos realizan inferencias materiales (Rosenberg, 2007) o inferencias protológicas (Bermúdez, 2003). Hasta donde sé, carecemos, en cambio, de una explicación análoga de los razonamientos inferenciales que involucran contenidos de objeto. Carecemos también de una explicación sobre cómo podrían interactuar entre sí estados con contenidos objetuales y estados con contenidos proposicionales, para dar lugar a inferencias teóricas o prácticas. De hecho, los mismos defensores de los contenidos de objeto reconocen que esta es un área a la cual es preciso redireccionar sus esfuerzos de modo urgente (Montague y Grzankowski, 2018).

4. Conclusión

En este trabajo, he procurado reconstruir el intento de Baron y colegas por identificar y describir un modo de interacción que puede tener lugar en criaturas carentes tanto de lenguaje como de pensamiento proposicional (en sentido pleno): la llamada “triangulación intermedia”. Quien participa de este tipo de vínculo triangular con otro semejante y con un objeto del entorno común no se limita a responder causalmente a ambos, sino que da respuestas afectivas y comunicativas, consistentes en la producción

de comportamientos expresivos naturales –como los aullidos, las carcajadas, el llanto, los gritos de alarma, los gestos de sorpresa, etc.–. Tales comportamientos, se nos dice, han sido naturalmente diseñados para mostrar los estados psicológicos de quien los realiza a otros que estén dotados cognitivamente para reconocerlos. En este sentido, pueden operar como genuinos precursores de nuestras formas más sofisticadas de comunicación lingüística.

Un punto central que los autores señalan recurrentemente es el siguiente: los estados mentales que estas criaturas manifiestan, mediante distintos comportamientos expresivos, no tienen por qué tener el tipo de complejidad y sofisticación de las actitudes proposicionales que los filósofos suelen adscribir a las criaturas humanas lingüísticamente competentes. De hecho, una de las razones que los diferencian de tales actitudes, se nos dice, es el carácter no proposicional de sus contenidos. Según vimos, sin embargo, hay razones en pugna que nos llevan a pensar que cabe adscribir a estos filósofos dos modos distintos de entender tal noción: como contenidos de objeto, o, alternativamente, como contenidos que representan estados de cosas, pero que, pese a ello, no llegan a satisfacer todos los requerimientos para ser considerados como contenidos proposicionales en un sentido pleno.

Ante esta disyuntiva, me pareció interesante intentar abordar el debate desde un ángulo diferente: el de examinar cuál posición resultaría más adecuada, dados sus intereses explicativos, a la hora de dar cuenta del tipo de contenidos mentales que se manifiestan mediante comportamientos expresivos. Con este nuevo propósito en mira, esbocé brevemente tres vías alternativas que se podrían seguir: a) entender la totalidad de los contenidos que se expresan mediante comportamientos expresivos como contenidos de objeto; b) entenderlos como

contenidos que representan estados de cosas, pero que no llegan a ser contenidos proposicionales plenos; y c) adoptar un enfoque híbrido, de acuerdo con el cual los comportamientos expresivos pueden manifestar, en distintas ocasiones, tanto estados mentales con contenidos de objeto, como contenidos no proposicionales que versen sobre estados de cosas. Esta opción, sospecho, es la más adecuada si se quiere capturar el amplio espectro de conductas y capacidades cognitivas presentes en distintas criaturas sin lenguaje. Una justificación más elaborada de tal tesis debe ser dejada, sin embargo, como tarea pendiente para un futuro trabajo.

Referencias bibliográficas

- Bar-on, D. y Priselac. M. (2011). "Triangulation and the beasts". En G. Preyer y C. Amoretti (Eds.), *Triangulation: From an Epistemological Point of View* (pp 121-152). Frankfurt: Ontos Verlag.
- Bar-on, D. (2010). "Expressing as 'Showing What is Within': On Mitchell Green's Self-Expression". *Philosophical Books*, 51(4), pp. 212-227.
- Bar-on, D. (2013). "Expressive Communications and Continuity Skepticism". *Journal of Philosophy* 110(6), pp. 293-330.
- Bar-on, D. (2018). "Minding the Gap: In Defense of Mind-mind Continuity". En K. M. Cahill y T. Raleigh. (Eds.), *Wittgenstein and Naturalism* (pp.177-203). New York: Routledge.

- Bar-on, D. y Green, M. (2010). "Lionspeak: Communication, Expression and Meaning". En J. O'Shea y E. Rubenstein (Eds.), *Self, Language and World: Problems from Kant, Sellars and Rosenberg* (pp. 89-106). Atascadero, CA: Ridgeview Publishing Company.
- Bermúdez, J.L. (2003). *Thinking without Words*, New York: Oxford University Press.
- Burge, T. (2010a). "Steps towards Origins of Propositional Thought". *Disputatio*, IV(29), pp. 39-67.
- Burge, T. (2010b). *Origins of Objectivity*. New York: Oxford University Press.
- Camp, E. (2015). "Logical Concepts and Associative Characterizations". En E. Margolis y S. Laurence (Eds.), *The Conceptual Mind: New Directions in the Study of Concepts* (pp. 591-621). Cambridge, MA: MIT Press.
- Camp, E. (2004). "The Generality Constraint and Categorical Restrictions". *Philosophical Quarterly* 64 (2015), pp. 209-231.
- Crane, T. (2009). "Is Perception a Propositional Attitude?". *Philosophical Quarterly*, 59(236), pp. 452-469.
- Crane, T. (2013). *The Objects of Thought*. Oxford: Oxford University Press.
- Cussins, A. (1992). "Content, Embodiment and Objectivity: The Theory of Cognitive Trails". *Mind*, 101(404), pp. 651-688.
- Danón, L. (2016). "Opacidad referencial y atribución intencional a animales sin lenguaje". *Principia* 20(2), pp. 143-164.
- Davidson, D. (1975). "Pensamiento y habla". En *De la verdad y la interpretación* (pp. 164-168). Barcelona: Gedisa.
- Davidson, D. (1982). "Rational Animals". *Dialectica*, 36(4), pp. 317-327.

- Davidson, D. (1999). The Emergence of Thought. *Erkenntnis*, 51(1), 7-17.
- Davidson, D. (2001a). "Externalisms". En P. Kotatko, P. Pagin y G. Segal (Eds.), *Interpreting Davidson; Selected Papers from the 1996 Karlovy Vary Symposium in Analytic Philosophy* (pp. 5-13). Palo Alto: CSLI Press.
- Davidson, D. (2001b). "What Thought Requires". En *Problems of Rationality*. Nueva York: Oxford University Press [2004].
- De Vries, W. (1995). "Sellars, Animals and Thoughts". En *Problems of Wilfrid Sellars*. Recuperado de <https://goo.gl/mSMcsg>.
- Evans, G. (1982). *The Varieties of Reference*. Nueva York: Oxford University Press.
- Glock, H. J. (en prensa). Animal minds: conceptual problems.
- Gluer, K. (2008). "Triangulation". E. LePore y B. Smith (Eds.), *The Oxford Handbook of Philosophy of Language*. New York: Oxford University Press.
- Grzankowski, A. (2013). Non-Propositional Attitudes, *Philosophy Compass*, 8(12), pp. 1123-1137.
- Grzankowski, A. (2015). "Not All Attitudes Are Propositional". *European Journal of Philosophy* 23(3), pp. 374-391.
- Grzankowski, A. (2016). "Attitudes Towards Objects". *Nous*, 50, pp. 314-328.
- Ludwig, K. (2011). "Triangulation triangulated". Amoretti, C. y G. Preyer (Eds.), *Triangulation: From an Epistemological Point of View* (pp. 69-96). Frankfurt: Ontos Verlag.
- Mandelbaum, E. (2013). "Against alief". *Philosophical Studies* 165(1), pp. 197-211.
- Montague, M. (2007). "Against Propositionalism". *Nous*, 41(3), pp. 503-518.

- Montague, M. y Grzankowski, A. (2018). "Non-propositional intentionality: An introduction". En M. Montague y A. Grzankowski (Eds.), *Non-Propositional Intentionality* (pp. 1-18). Nueva York: Oxford University Press.
- Pagin, P. (2001). "Semantic triangulation". En P. Kotatko, P. Pagin y G. Segal (Eds.), *Interpreting Davidson* (pp. 199-212), Stanford: CSLI Publications.
- Rosenberg, J. (2007). *Wilfrid Sellars: Fusing the Images*. Oxford: Oxford University Press.
- Schroeder, T. (2006). "Reply to Critics". *Dialogue*, 45(1), pp. 165-174.
- Searle, J. (1983). *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Sellars, W. (1969). "Language as Thought and as Communication". *Philosophy and Phenomenological Research* XXIX, 4, pp. 506-527.
- Sellars, W. (1981). "Mental Events". *Philosophical Studies* 39, pp. 325-345.
- Verheggen, C. (2007). "Triangulating with Davidson". *The Philosophical Quarterly*, 57(226), pp. 96-103.

Algunas reflexiones en torno al conocimiento de las otras mentes

DANIEL KALPOKAS (UNC – CONICET)

1. Introducción

¿Cómo es que llegamos a conocer los estados mentales ajenos? ¿Qué relación existe entre las conductas expresivas, por un lado, y los estados mentales expresados por ellas, por otro? ¿Y qué relación existe entre las conductas expresivas, por un lado, y nuestro conocimiento de las otras mentes, por otro? Es obvio que estas tres preguntas están estrechamente conectadas, al menos si bien nos hacemos la primera de las preguntas mencionadas. Por ejemplo, si concebimos las conductas expresivas como constitutivas, al menos en parte, de los estados mentales mismos, entonces podemos sostener plausiblemente que los estados mentales ajenos pueden ser conocidos, en ciertas ocasiones, mediante observación directa. En cambio, si pensamos que las conductas expresivas son meros efectos de los estados mentales que las causan y, por ende, no creemos que sean ingredientes constitutivos de los estados mentales mismos, parece que tendremos que reconocer que los estados mentales ajenos resultan inobservables para nosotros, y, consecuentemente, que llegamos a cono-
cerlos, cuando así lo hacemos, solo inferencialmente.

En este trabajo, quiero focalizarme, principalmente, en la primera de las preguntas mencionadas, aunque una respuesta (tentativa, por cierto) a ella implique inevitablemente algunos compromisos que se vinculan con respuestas a las otras preguntas. Tiene que quedar claro desde el comienzo que, al concentrarme en la pregunta sobre cómo es que conocemos los estados mentales ajenos, me sitúo en el terreno epistemológico y, por ende, en el nivel personal de análisis de los procesos o procedimientos involucrados en el conocimiento de lo que sucede en las mentes de las otras personas. Mi pregunta es, pues, ¿cómo es que, al menos en ocasiones, sabemos que otras personas están en tal o cual estado mental?

2. Estrategias inferencial y no inferencial

En la literatura sobre el tema, hay, desde luego, varias alternativas. Con todo, si las consideramos en vistas del modo en que, de acuerdo con ellas, llegamos a conocer las mentes ajenas, podemos reducirlas, básicamente, a dos grandes estrategias: la inferencial y la no inferencial (McNeill, 2012). La estrategia inferencial sostiene lo siguiente: al conocer los estados mentales ajenos, partimos de la observación directa de sus conductas expresivas (esto es, gestos y otros tipos de comportamiento) y, a partir de esos datos observacionales, mediante un proceso inferencial de algún tipo, llegamos a atribuir o postular ciertos estados mentales como causa de las conductas expresivas que han servido de base evidencial para llevar a cabo la mencionada inferencia. La estrategia inferencial está paradigmáticamente ejemplificada por dos teorías de la atribución mental bien conocidas: la teoría de la teoría (de aquí en adelante, TT) y la teoría de la simulación (de aquí en

adelante, TS). Según la primera, a partir de la observación de la conducta expresiva de los otros, elaboraríamos una teoría acerca de los estados mentales que causan la conducta observada. Según la segunda, en cambio, la observación de la conducta de las otras personas permitiría iniciar un proceso de simulación de los estados mentales ajenos, causantes de la mencionada conducta, en virtud del cual atribuiríamos ciertos estados mentales.

En contraste, la estrategia no inferencial sostiene típicamente que, al menos para el caso de ciertos estados mentales (v.g. enojo, alegría, tristeza, etc.), llegamos a conocer los estados mentales ajenos mediante observación directa. La idea, aquí, es que la percepción directa de la conducta expresiva de los otros provee, por sí misma, conocimiento de sus estados mentales (de aquí en adelante, PD). Aunque no es preciso restringir la estrategia no inferencial al caso de la percepción, este parece ser el camino que se da por descontado en las discusiones sobre el tema.

3. Comunicación de los estados mentales

Ahora bien, pienso que estas alternativas teóricas, esto es, la TT, la TS y la PD, no solo dejan un espacio interesante para matices, clarificaciones y variedades dentro de cada una de ellas, sino que, además, no agotan las posibilidades acerca de cómo es que llegamos a saber de los estados mentales de los otros. Retomemos el problema de cómo conocemos, al menos en ciertas ocasiones, las otras mentes. ¿Es verdad que nuestro conocimiento de las otras mentes es alcanzado, en todos los casos, tal como la TT o la TS lo sugieren, por un lado, o bien como los defensores de la PD sostienen, por el otro? Me parece bastante obvio

que esta disyunción no es exhaustiva. Por ejemplo, muchas veces sabemos de los estados mentales ajenos porque, simplemente, los otros nos dicen que se encuentran en tal o cual estado mental. En tales casos, un acto de habla, realizado con la intención de comunicar el estado mental en el que se encuentra la persona, es el medio por el cual, si la persona es sincera, llegamos a saber de su mente. “Estoy cansado”, “Estoy deprimido”, “Estoy contento”, “Estoy muy enojado” son actos de habla expresivos que nos informan, típicamente, de lo que está sucediendo en la mente del hablante⁴⁰. En otras ocasiones, sin embargo, puede haber intención de comunicar un estado mental sin recurrir al lenguaje. En tales casos, el sujeto apela a ciertos gestos, que incluso pueden tener un uso convencional, para comunicar a otro que se halla en tal o cual estado mental. En contraposición al primer caso, en el que un acto de habla, que se vale de expresiones lingüísticas, provee la información sobre el estado mental del hablante, llamemos a este segundo caso “acto gestual”, para dar a entender que el sujeto efectúa ciertos gestos con el fin de comunicar a otro cuál es el estado mental en el que se encuentra. Como en el caso en el que un hablante nos dice lingüísticamente que se halla en tal o cual estado mental, en el caso en el que una persona nos comunica, mediante ciertos gestos, que le duele la cabeza, por ejemplo, o que está nerviosa, también podemos llegar a saber, si la persona en cuestión es sincera, que ella se encuentra en el estado mental en el que dice estar. Lo que tienen en común ambos tipos de actos, el acto de habla y el gestual, es que son actos que se realizan con la intención de comunicar a otro el propio

⁴⁰ Este es un punto que no se le ha escapado a Austin. Véase Austin (1961, 1975).

estado mental, y, por ende, el significado de lo expresado es, en ambos casos, del orden de lo que Grice llamó “significado no natural” (Grice, 1991).

Cuando reflexionamos sobre esta posibilidad (esto es, la posibilidad de saber que una persona está en tal o cual estado mental en virtud de lo que nos comunica, verbal o no verbalmente), resulta bastante claro que las opciones sobre cómo llegamos a conocer los estados mentales ajenos, presentadas antes, no resultan exhaustivas. En efecto, cuando una persona sincera nos dice, o nos da a entender mediante gestos, que tiene dolor de cabeza, por ejemplo, y nosotros aceptamos lo que nos comunica, estamos en posición de saber que la persona en cuestión tiene dolor de cabeza. Pero no llegamos a estar en posesión de este conocimiento, ni en virtud de un proceso inferencial como el postulado por la TT o la TS, ni en virtud de que percibimos directamente su estado mental. Simplemente, la persona nos comunica que tiene dolor de cabeza. Así, podría decirse que este es un conocimiento al que arribamos por una vía alternativa a las ya mencionadas, a saber, el testimonio directo de la persona en cuestión. (Desde luego, puede ser el caso que sepamos que a Juan le duele la cabeza porque Pedro nos lo ha dicho. Para simplificar, solo considero los casos en los que el sujeto cuyos estados mentales conocemos es el que nos brinda el testimonio). Cómo hemos de entender el conocimiento por testimonio es, desde luego, un problema en sí mismo. No pretendo adentrarme aquí en las discusiones propias de la epistemología del testimonio⁴¹. Simplemente quiero señalar algo que muchas veces se pasa por alto, o que no se tiene suficientemente en cuenta, en las discusiones acerca del conocimiento de las otras

⁴¹ Para una visión general sobre los problemas involucrados en la epistemología del testimonio y, en particular, sobre la cuestión de si el testimonio es una forma inferencial o no inferencial de justificación y conocimiento, véase Adler (2012).

mentes: que, además de las debidas inferencias, como las postuladas por la TT y la TS, y de la percepción, el testimonio muchas veces constituye una fuente de conocimiento de las otras mentes.

4. Un argumento en favor de la estrategia perceptiva

Ahora bien, ¿cómo es que llegamos a conocer los estados mentales de las otras personas cuando estas no nos dicen, verbal o no verbalmente, en qué estado mental se encuentran? Ante la incapacidad de conocer los estados mentales ajenos telepáticamente, no tenemos otro remedio que atender a las conductas expresivas (gestos y demás comportamientos) de los otros. Es aquí en donde encuentran su plena aplicación las alternativas teóricas ya mencionadas, como la TT, la TS y la PD. Sin pretender negar que, quizás en muchas ocasiones, tal vez conozcamos los estados mentales ajenos por medio de procesos tales como los postulados por la TT y la TS, quisiera hacer aquí algunas consideraciones exploratorias sobre la idea de que, al menos en ciertas ocasiones y para ciertas clases de estados mentales, podemos percibir, en un sentido a especificar, los estados mentales ajenos.

La idea de que, al menos en ciertas ocasiones, percibimos los estados mentales ajenos es actualmente el foco de un encendido debate⁴². Sin embargo, pese a los distintos desarrollos propuestos, resulta difícil encontrar en la literatura sobre el tema argumentos de por qué habríamos

⁴² Véase Krueger (2012); McNeill (2012); Krueger y Overgaard (2012); Cassam (2007); Smith (2010); McGeer (2009); Gallagher (2008); Zahavi (2011); Stout (2011); Danón y Kalpokas (2017). Dentro de la tradición analítica, los antecedentes de esta idea remiten a Wittgenstein (1953, 2009); McDowell (1998) y Dretske (1973).

de aceptar la tesis de esta idea. Típicamente, los distintos autores que están de acuerdo con ella dan por supuesta su plausibilidad y, a renglón seguido, tratan de explicarla apelando a algún modelo perceptivo. Aunque yo también diré algo acerca de lo que considero que es percibir algunos de los estados mentales ajenos, quisiera comenzar presentando un argumento a favor de dicha tesis.

El argumento apela a la fenomenología de la experiencia perceptiva. Si, al observar el comportamiento de una persona, esta se nos aparece visualmente, por ejemplo, como si estuviera en tal o cual estado mental aun cuando sepamos que, en realidad, no está en aquel estado mental y, por tanto, aun cuando no estemos dispuestos a inferir, a partir de la observación de su conducta, que está en tal estado mental, entonces puede afirmarse que, fenomenológicamente hablando, existe para nosotros la posibilidad de experimentar (visualmente) el comportamiento de una persona *como si* estuviera en el estado mental relevante. La idea es que, al observar las conductas expresivas de una persona, no solo podemos experimentarlas como meros movimientos de un cuerpo humano, sino como vehículo expresivo del estado mental correspondiente.

Por ejemplo, supóngase que un actor de teatro intenta transmitir a su audiencia la impresión de que el personaje que está representando está en un estado de dolor extremo. En tal caso, su conducta expresiva será acorde al estado mental que intenta transmitir. Sus gestos, gritos y comportamientos pueden ser tan vívidos que impresionen sobremanera, incluso, a su audiencia. Aunque esta sepa en el fondo que se trata de una actuación y, por tanto, no esté dispuesta a inferir, de la conducta observada, la conclusión de que el actor se encuentra realmente en ese estado mental, experimentará, sin duda, la conducta expresiva del actor como conducta de dolor extremo. El

ejemplo muestra, creo, que tenemos la capacidad de experimentar las conductas expresivas ajenas como expresiones de sus estados mentales, incluso cuando no creamos, o no tengamos la disposición de inferir, que la persona en cuestión se encuentra en el estado mental indicado por su comportamiento⁴³.

5. Percepción de estados mentales ajenos

¿Cómo hemos de entender la idea de que podemos percibir al menos algunos de los estados mentales de las otras personas? Hay, desde luego, distintas maneras de entenderla⁴⁴; pero quisiera comenzar sugiriendo una forma de entender las conductas expresivas mismas. Se podría pensar que las conductas expresivas que no son ejecutadas por el agente con la intención de comunicar sus estados mentales significan naturalmente los estados mentales en los que se encuentra el agente. De acuerdo con la noción de significado natural tal como se viene utilizando desde Grice (1991), cuando *x* (cierta conducta, podemos decir aquí) significa naturalmente *p* (la presencia de cierto estado mental), *x* implica *p*. Dicho en otros términos, el significado natural es factivo. Las correlaciones entre *x* y *p* no deben ser accidentales. La conexión o dependencia entre *x* y *p* debe existir en la naturaleza. De acuerdo con esta idea,

⁴³ El ejemplo puede alterarse ligeramente para mostrar el mismo punto. Podemos suponer ahora que una persona que da muestras de dolor extremo está realmente sufriendo. Sin embargo, supongamos que se nos ha hecho creer que se trata de una muy buena actuación y que, por tanto, la persona que estamos observando no está sufriendo realmente lo que parece sufrir. Nuevamente, como en el caso anterior, aunque al observar su conducta no tengamos la disposición de inferir que está sufriendo, no obstante, nos parecerá, fenomenológicamente hablando, que está sufriendo cierto tipo de dolor.

⁴⁴ Véase nota 3.

pues, cuando tiene lugar cierta conducta expresiva, esta significa naturalmente cierto estado mental del agente. Si la conducta *x* realmente significa naturalmente el estado mental *p*, entonces el agente está, en efecto, en el estado mental *p*. Por supuesto, alguien puede tomar, equivocadamente, a cierta conducta *y* como significante del estado mental *q*. Pero, en tal caso, no debemos decir que *y* significa naturalmente *q*⁴⁵. Así, hablar de conductas expresivas es hablar de conductas que poseen significado natural. Tales conductas significan naturalmente el estado mental que es expresado por ellas.

La idea de significado natural, aplicada a las conductas expresivas de los agentes, permite arrojar cierta luz a la idea según la cual percibimos, en cierto sentido, los estados mentales ajenos. Al observar las conductas expresivas de los otros, podemos percibirlas *como* expresiones de tales y cuales estados mentales porque podemos percibirlas *como* significantes naturales de tales y cuales estados mentales. Según este punto de vista, resulta simplista y, tal vez, desorientador decir que literalmente vemos el dolor de una persona en su rostro, pues, intuitivamente hablando, no parece que veamos el dolor en su rostro en el mismo sentido en que podemos ver su nariz, sus ojos o sus cejas. Antes bien, reconocemos perceptivamente que ciertos gestos y comportamientos, que aparecen típicamente cuando las personas sienten dolor, o que aparecen típicamente al menos en la persona en cuestión, significan naturalmente un estado de dolor. La idea de significado natural es introducida aquí para tratar de aclarar cómo es que, en ocasiones, logramos percibir ciertas conductas *como* expresiones de ciertos estados mentales. Lo que un sujeto que es capaz de percibir estados mentales ajenos

⁴⁵ Esto vale incluso en el caso de que contingentemente se dé *q*.

tiene que saber es que hay una correlación no contingente entre el hecho de que la persona observada lleve a cabo ciertas conductas expresivas y el hecho de que tal persona se encuentre en el estado mental correspondiente. Así, al poseer el conocimiento de que ciertas conductas expresivas y patrones conductuales significan naturalmente tales o cuales estados mentales, dichas conductas y patrones pueden presentársenos, fenomenológicamente hablando, como expresando los estados mentales que, si estamos en lo correcto, efectivamente son. No se trata, pues, de que percibimos ciertos movimientos corporales y, a partir de ello, inferimos (cualquiera sea la naturaleza de esta inferencia) la presencia del estado mental correspondiente. Antes bien, el conocimiento de que tales y cuales conductas expresivas significan naturalmente tales y cuales estados mentales introduce una diferencia en la fenomenología de nuestra experiencia: nos permite percibir las conductas expresivas en cuestión como expresiones, o significantes naturales, de la presencia de los estados mentales correspondientes⁴⁶.

Ahora bien, se ha señalado, correctamente a mi juicio, que la percepción es una función de las expectativas que nos formamos cuando percibimos algo⁴⁷. La idea es que el contenido perceptivo no está constituido solamente por el modo en que un objeto, desde cierta perspectiva y en determinadas condiciones ambientales, afecta los sentidos, sino, también, por las expectativas que tenemos respecto al modo en que ese mismo objeto afectará nuestros sentidos desde otras perspectivas y condiciones

⁴⁶ Esta tesis supone, desde luego, la plausibilidad de la penetrabilidad cognitiva que ciertos tipos de conocimientos y habilidades tienen sobre la experiencia perceptiva. Al respecto, véase, por ejemplo, Siegel (2006, 2010).

⁴⁷ Véase Husserl (1970, 1982); Sellars (1978); Searle (1983); Noë (2004); y Kalpokas (2017).

ambientales. Por ejemplo, al ver un tomate de frente, esperamos que, de movernos de tal y cual modo, o de moverse el tomate mismo de tal o cual modo, el tomate aparezca a nuestros ojos de tal y cual forma. Asimismo, al oír cierto sonido en la calle, que identificamos inicialmente como el sonido de un automóvil acercándose, esperamos que su intensidad se incremente a medida que pasan los segundos. La experiencia perceptiva incluye, pues, lo directamente presentado (la cara visible del tomate, el sonido del auto oído en cierto momento) y aquello que es amodalmente copresentado (la cara de atrás del tomate, el incremento del sonido del auto a medida que se acerca)⁴⁸. La idea no es, pues, que al percibir cierta parte o aspecto de un objeto inferimos sus otras partes o aspectos. La idea es, más bien, que los objetos que percibimos desde cierta perspectiva y en determinadas condiciones ambientales se nos aparecen, fenomenológicamente hablando, como si tuvieran otras partes o aspectos que, por el momento, no son directamente percibidos.

Algo semejante es de esperar que suceda con la percepción de las conductas expresivas de las otras personas. Al ver, por ejemplo, ciertos gestos que inicialmente identificamos como gestos de enojo, esperamos que el comportamiento de la persona en cuestión siga un patrón consistente con lo que hemos identificado perceptivamente como un gesto de enojo. Las conductas expresivas y comportamientos de las personas constituyen un cierto tipo de eventos o procesos. Como sucede con cualquier otra clase de eventos o procesos, estas conductas expresivas tienen que revelar ciertos patrones para poder ser identificadas como las conductas que son. En consecuencia, si, después de haber percibido un gesto que inicialmente

⁴⁸ Tomo aquí la terminología de lo presentado vs. lo copresentado de Smith (2010).

fue visto como un gesto de enojo, la persona del gesto se comporta de un modo que resulta completamente inconsistente con lo que creíamos haber identificado como un comportamiento de enojo, tendremos que albergar la sospecha de que tal vez lo que percibimos no fue, después de todo, un gesto de enojo. En tal caso, decimos que nuestras expectativas se han visto defraudadas: lo copresentado amodalmente en nuestra experiencia no se condice con lo que pensamos que era cuando vimos el gesto presentado. Por tanto, tenemos que decir que lo directamente presentado en la experiencia (el gesto) no significa naturalmente, en verdad, enojo, sino que, más bien, *parece* significar un gesto de enojo. En cambio, si el patrón conductual de la persona en cuestión es consistente con lo que entendemos que es un estado de enojo, nuestras expectativas al ver el gesto como un gesto de tal emoción no se verán defraudadas: lo que en un comienzo identificamos perceptivamente como un gesto de enojo resulta consistente con el desenvolvimiento de los hechos ulteriores. Podemos decir, pues, que estamos viendo, o que hemos visto, a la persona enojada en el sentido de que estamos viendo, o hemos visto, sus conductas expresivas como significantes naturales de enojo.

Así, de acuerdo con el enfoque propuesto aquí, podemos percibir, al menos en ciertas situaciones, los estados mentales ajenos, en el sentido de percibir sus gestos y conductas expresivas como significantes naturales de los estados mentales que tales conductas expresan. Las conductas expresivas de los otros se nos aparecen, fenomenológicamente hablando, *como* que son conductas de dolor, de enojo, de alegría, etc., en virtud del significado que les atribuimos. Para ello, se requiere, desde luego: a) estar en posesión de cierto tipo de conocimiento de trasfondo acerca de los estados mentales en general y del modo en

que pueden aparecerse sensorialmente en las correspondientes conductas expresivas; y b) aceptar la tesis de que tal conocimiento es capaz de afectar la fenomenología de nuestra experiencia perceptiva. Finalmente, se ha sugerido que c) la percepción es esencialmente perspectivística; esto quiere decir que percibir un objeto, propiedad, evento o proceso completos requiere de la percepción amodal de los aspectos, o partes, del ítem percibido que, por el momento, no afectan nuestro aparato perceptivo. En el caso particular de las conductas expresivas, percibirlas como significantes de tal o cual estado mental requiere de la percepción amodal de un patrón conductual consistente con el gesto o conducta que inicialmente ha sido perceptivamente identificada como expresión de tal o cual estado mental.

6. Posibles objeciones

Veamos algunas objeciones que pueden hacerse al punto de vista sugerido aquí. Tratar de responderlas puede ayudarnos a comprender mejor la sugerencia en cuestión.

En primer lugar, podría objetarse que, al apelar a la idea de significado natural, estoy pasando por alto la distinción wittgensteiniana entre criterios y síntomas (Wittgenstein, 1935; 1994, p. 53). Los criterios sirven para definir un estado, proceso o propiedad. Por ejemplo, si se pregunta a un médico por qué dice que un paciente tiene angina, su respuesta "Porque he encontrado en él tal y cual bacilo en su sangre" proporciona el criterio definidor de la angina. Un síntoma, en cambio, es algo que la experiencia nos ha enseñado que coincide, de un modo u otro, con el fenómeno que es nuestro criterio definidor. Así, si a la pregunta antedicha sobre por qué el médico dice que su

paciente tiene angina se responde con “Porque tiene la garganta inflamada”, entonces se ha dado un síntoma de la angina. Decir, por tanto, que un hombre tiene angina si se encuentra en su sangre el bacilo cuya presencia define qué es la angina, es incurrir en una tautología. En cambio, decir que un hombre tiene angina porque tiene la garganta inflamada es hacer una hipótesis que, llegado el caso, podría resultar falsa.

Sin embargo, la distinción de Wittgenstein no captura exactamente la idea de significación natural de Grice. Cuando decimos que *x* significa naturalmente *p*, no estamos expresando meramente una convención o una definición (como es el caso de la noción wittgensteiniana de criterio), por un lado, ni estamos refiriéndonos a una regularidad contingente (como es el caso de la noción de síntoma), por el otro. Cuando decimos que ciertas manchas significan naturalmente sarampión⁴⁹, no estamos indicando que las manchas en cuestión son meros síntomas de sarampión (en cuyo caso el paciente podría no tener sarampión, después de todo), ni estamos dando una definición de lo que llamamos “sarampión”; antes bien, estamos señalando una correlación no derrotable, por así decirlo, entre dos hechos, esto es, tener ciertas manchas en la piel y tener sarampión. Es cierto que uno podría tomar esta correlación para usarla como un criterio definidor, si así quisiera. Pero la idea de significado natural indica algo anterior, a saber, que existe la correlación antedicha en la naturaleza misma. En consecuencia, y de modo semejante, decir que cierta conducta expresiva significa naturalmente enojo, digamos, no es definir lo que nosotros llamamos “enojo”, ni indicar un síntoma de lo que probablemente sea, en

⁴⁹ Este es uno de los ejemplos del propio Grice. Véase Grice (1991, p. 213).

la persona en cuestión, un estado de enojo. Es, más bien, señalar cierta correlación natural entre la conducta expresiva y el estado de enojo.

En segundo lugar, se podría atacar mi versión de PD diciendo, como una vez dijo Austin (1961/1975, pp. 108-111)⁵⁰, que hablar de signos, en referencia a la conducta expresiva, sugiere un contraste entre lo que efectivamente podemos percibir (esto es, los signos) y aquello que es significado por lo que percibimos (es decir, el estado mental). Así, podría argumentarse que introducir aquí la idea de signo nos aleja de la idea de que podemos percibir, al menos en ocasiones, los estados mentales ajenos: percibimos los signos externos de los estados mentales, no los estados mentales mismos.

Sin duda, al responder a esta objeción, uno puede apelar a la distinción de Dretske (1969) entre ver primaria y secundariamente. Ver primariamente es ver directamente el hecho en cuestión. En cambio, ver secundariamente es ver cierto hecho en virtud de que podemos ver otro hecho conectado apropiadamente con el primero. Así, podríamos decir que vemos *primariamente* ciertas conductas expresivas y que, al hacer esto, logramos ver *secundariamente* el estado mental apropiadamente vinculado a ellas. Pero tengo en mente una respuesta a la objeción antedicha que, me parece, es más radical. Es esta: no hay nada en el concepto de significado natural que nos obligue a pensar que un signo natural no pueda ser parte del estado o proceso mayor que es significado por él. Podemos decir, por ejemplo, que el color amarillo de un limón significa naturalmente cierto estado suyo -su madurez- sin implicar necesariamente con esto que su color sea algo completamente

⁵⁰ En verdad, Austin hace la apreciación que estoy reconstruyendo respecto de los signos porque defiende la idea de que los estados mentales, tales como el enojo, pueden, en realidad, observarse directamente.

diferente de ese estado. Podemos decir, así, que el color amarillo es una *manifestación*, entre otras, de su madurez, o que es un *componente* de lo que llamamos madurez de un limón. Similarmente, podemos decir que ciertos gestos –el ceño fruncido, ciertos ademanes, etc.–, en la ocasión apropiada⁵¹, no solo significan naturalmente enojo, sino que, además, son componentes de ese estado mental mismo⁵². En tal sentido, podemos decir que las conductas expresivas relevantes vehiculan, al menos en parte, el enojo, y, así, lo manifiestan. La idea no es sugerir que el estado que caracterizamos como de enojo se agota en sus conductas expresivas; más bien, la idea es que las conductas expresivas constituyen un componente de dicho estado mental, el cual típicamente involucra, además de las conductas expresivas, otros ingredientes, como, por ejemplo, el sentimiento de enojo.

Una tercera objeción podría rezar así: podría suceder, y con cierta frecuencia sucede de hecho, que las conductas expresivas no vayan acompañadas por los debidos procesos internos (sentimientos, sensaciones, etc.) que caracterizan al estado mental en cuestión. Alguien puede hacer gestos de dolor, pongamos por caso, o de enojo, sin tener dolor o estar enojado. Por tanto, no hay una correlación estricta entre las conductas expresivas y los estados mentales que se supone que son expresados por ellas y, debido a esto, tampoco puede afirmarse que las conductas expresivas sean componentes de los estados mentales correspondientes. Como corolario de ello, también se podría señalar

⁵¹ Con respecto a la relevancia que poseen las ocasiones específicas en las que tienen lugar los estados mentales que llegamos a conocer, véase Austin (1961/1975, p. 112).

⁵² Para la idea según la cual las conductas expresivas pueden ser concebidas como partes constitutivas de los estados mentales mismos, véase, por ejemplo, Krueger y Overgaard (2012), y Danón y Kalpokas (2017).

que, aunque ciertamente podemos percibir las conductas expresivas, no podemos hacer esto con respecto a los estados mentales relevantes.

Esta objeción puede responderse del siguiente modo: es cierto que puede suceder, y a veces sucede de hecho, que ciertas conductas expresivas típicas no vayan acompañadas por sus correspondientes estados internos; sin embargo, en tales casos tenemos que abstenernos de decir que dichas conductas expresivas significan naturalmente sus estados mentales correspondientes. Más bien, tenemos que decir que *parecen* significarlos naturalmente. Consecuentemente, tenemos que decir también que solo *parecen* ser componentes de cierto estado mental, aunque en realidad no lo sean. Esto revela la consistencia del presente enfoque con una teoría disyuntiva de la experiencia. Haciendo caso omiso de sus distintas variedades⁵³, se puede decir que el disyuntivismo se caracteriza por sostener que las experiencias sensoriales, cuando son un caso de percepción, son de una clase completamente diferente de las experiencias alucinatorias/ilusorias, incluso si ambas experiencias resultan indistinguibles para el sujeto de experiencia. No hay, pues, una única clase de estado mental experiencial que caracterice a ambos tipos de experiencia. Así, conectando el disyuntivismo perceptivo con el enfoque que estoy presentando aquí, se puede afirmar que, al percibir ciertas conductas expresivas, puede suceder que estas signifiquen naturalmente cierto estado mental, o bien que solo parezcan significarlo. En el primer caso, percibimos las conductas expresivas como componentes del estado mental correspondiente; en el segundo, en cambio, solo parece que las estamos percibiendo como componentes de tal estado. El punto crucial es este:

⁵³ Sobre las distintas variedades del disyuntivismo, véanse los artículos compilados en Byrney Logue (2009) y en HaddockyMacpherson (2008).

en ambos casos experimentamos las conductas expresivas como significantes de algo más que de ellas mismas, esto es, como expresiones no meramente de ciertos movimientos y gestos del agente, sino también de su estado mental. Cuando estamos en lo correcto, podemos adquirir conocimiento perceptivo de los estados mentales ajenos. Existe, pues, cierta prioridad conceptual de la experiencia verídica por sobre la ilusoria: no nosaría *parecer* que ciertas conductas expresivas signifiquen naturalmente tal y cual estado mental si no fuera porque, al menos en ocasiones, conductas semejantes significan *realmente*, de un modo natural, tal y cual estado mental.

Por último, se podría hacer la objeción opuesta a la recién tratada: supóngase que existe un aprendiz de estoico, esto es, una persona que, aunque tiene estados mentales de dolor, enojo, tristeza, alegría, euforia, etc., nunca da señales de los estados mentales en los que se encuentra. Se trata de una persona que, podemos imaginar, ha aprendido a dominar sus pasiones a tal punto que, aunque posea los sentimientos y sensaciones que todos tenemos, nunca los exterioriza ante nadie. Con respecto a esta persona, no habría conducta expresiva que nos pudiera guiar con respecto a sus estados mentales y, por ende, no podríamos afirmar que podemos percibirlos directamente.

Desde luego, en un caso así, no podríamos hablar de percibir su conducta expresiva como significantes naturales de tal o cual estado mental. Con todo, hay que notar que tampoco podríamos decir que conocemos sus estados mentales inferencialmente, pues no habría una base evidencial a partir de la cual construir una inferencia (ya sea por medio de mecanismos como los postulados por la TT o como los postulados por la TS) cuya conclusión fuera un estado mental que pudiéramos atribuir a tal persona. En verdad, en un caso como el imaginado tendríamos

dificultades para entender a un agente de tales características como poseedor de estados mentales en general. A menos que nos hablara y nos dijera en qué estados mentales se encuentra, supongo que podríamos hacer, a lo sumo, conjeturas sobre algunos de sus propósitos y creencias observando cómo se desplaza en el mundo. Sin embargo, un caso límite como este plantearía un desafío para *todas* las teorías de la atribución mental, no solo para la propuesta aquí.

7. Conclusiones

En este trabajo, he recordado, en primer lugar, la importancia que posee el testimonio para el conocimiento de las mentes ajenas y la incidencia que la epistemología del testimonio puede tener en este ámbito del conocimiento; y he explorado, en segundo lugar, cómo la idea de que percibimos los estados mentales ajenos puede ser elaborada cuando la combinamos con la idea de que las conductas expresivas significan naturalmente los estados mentales significados por ellas. La articulación con la idea de significado natural provee una explicación interesante, a mi juicio, del sentido en el que puede decirse que percibimos las conductas expresivas de los otros como manifestaciones de sus estados mentales.

Referencias bibliográficas

- Adler, J., (2012). "Epistemological Problems of Testimony". *Stanford Encyclopedia of Philosophy*, <https://goo.gl/SB5Xco>.

- Austin, J., (1961/1975). "Otras mentes". En *Ensayos filosóficos* (pp. 87-117). Traducido por A. García Suárez, Madrid: Revista de Occidente.
- Byrne, A. y Logue, H. (Eds.). (2009). *Disjunctivism*. Cambridge: The MIT Press.
- Cassam, Q. (2007). *The Possibility of Knowledge*. Oxford: Clarendon Press.
- Danón, L. y Kalpokas, D. (2017). "Perceiving mental states: Co-presence and constitution". *Filosofia Unisinos*, 18(2), pp. 87-97.
- Dretske, F. (1969). *Seeing and Knowing*. Chicago: University of Chicago Press.
- Dretske, F. (1973). "Perception and Other Minds". *Noûs*, 7, pp. 34-44.
- Gallagher, S. (2008). "Inference or interaction: social cognition without precursors". *Philosophical Explorations*, 11(3), pp. 163-174.
- Grice, P. (1991). "Meaning". En *Studies in the Way of Words* (pp. 213-223). Cambridge: Harvard University Press.
- Haddock, A. y Macpherson, F. (Eds.) (2008). *Disjunctivism: Perception, Action, Knowledge*. Oxford: Oxford University Press.
- Husserl, E. (1970). *The Crisis of European Sciences and Transcendental Phenomenology*. Trans. D. Carr. Evanston: Northwestern University Press.
- Husserl, E. (1982). *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy. First Book: General Introduction to a Pure Phenomenology*. Trans. F. Kersten, La Haya: Martinus Nijhoff.
- Kalpokas, D. (2017). "Experience and Justification: Revisiting McDowell's Empiricism". *Erkenntnis*, 82(4), pp. 715-738.

- Krueger, J., y Overgaard, S. (2012). "Seeing Subjectivity: Defending a Perceptual Account of Other Minds". En S. Miguens y G. Preyer (Eds.), *Consciousness and Subjectivity* (pp. 239-262). Heusenstamm: Ontos Verlag.
- Krueger, J. (2012). "Seeing mind in action". *Phenomenology and the Cognitive Sciences*, 11(2), pp. 149-173.
- McDowell, J. (1998). "Criteria, Defeasibility, and Knowledge". En *Meaning, Knowledge, y Reality* (pp. 369-394). Cambridge: Harvard University Press.
- McGeer, V. (2009). "The skill of perceiving persons". *The Modern Schoolman* 86 (2/3), pp. 289-318.
- McNeill, E.S. (2012). "Embodiment and the Perceptual Hypothesis". *The Philosophical Quarterly*, Vol. 62, (248), pp. 569-591.
- Noë, A. (2004). *Action in Perception*. Cambridge: The MIT Press.
- Searle, J. (1983). *Intentionality*. Cambridge: Cambridge University Press.
- Sellars, W. (1978). "The Role of Imagination in Kant's Theory of Experience". En Johnstone, H. Jr. (Ed), *Categories: A Colloquium*, Pensilvania: Pennsylvania State University.
- Siegel, S. (2006). "Which Properties Are Represented in Perception?". En T. Z. Gendler y J. Hawthorne (Eds.), *Perceptual Experience* (pp. 481-503). Oxford: Clarendon Press.
- Siegel, S. (2010). *The Contents of Visual Experience*. Oxford: Oxford University Press.
- Smith, J. (2010). "Seeing Other People". *Philosophy and Phenomenological Research*, 81(3), pp. 731-748.
- Stout, R. (2010). "Seeing the Anger in Someone's Face". *Proceedings of the Aristotelian Society Supplementary Volume*, 84, pp. 29-43.

- Wittgenstein, L., (1935/1994). *Los cuadernos azul y marrón.*
Traducido por F. Gracia Guillén, Barcelona: Planeta-Agostini.
- Wittgenstein, L. (1953/2009). *Philosophical Investigations.*
Traducido por G.E.M. Anscombe, P.M.S. Hacker Joachim Schulte. Hacker and Joachim Schulte, Singapur: Wiley-Blackwell.
- Zahavi, D. (2011). "Empathy and Direct Social Perception".
Review of Philosophy and Psychology, 2/3, pp. 541-558.

Percepción dinámica de expresiones faciales y atribución de emociones

ZOÉ SÁNCHEZ BARBIERI (UNC – CONICET)

1. Introducción

En muchas situaciones de la vida cotidiana, los humanos nos vemos inmersos en interacciones que requieren comprender el comportamiento de los individuos con los cuales interactuamos. Contamos, para ello, con diferentes herramientas cognitivas: en algunas ocasiones, razonamos acerca de las conductas pasadas del agente para entender su acción presente. En ciertas circunstancias, tendemos a generalizar, a partir de nuestro propio caso, para dar sentido a la conducta ajena, o bien comprendemos al otro aplicando un estereotipo (Andrews, 2012). En la mayoría de los casos, la comprensión, explicación y predicción del comportamiento ajeno involucra la detección y adscripción de una variedad de estados mentales, tales como creencias, deseos, intenciones, dolores, emociones, entre otros. Los psicólogos cognitivos y los filósofos de la mente han elaborado una serie de términos técnicos para dar cuenta de esta habilidad, entre los que se encuentran: “lectura de la mente”, “atribución intencional” y “psicología *folk*”. Específicamente, estas expresiones se utilizan para caracterizar, a grandes rasgos, el conjunto de habilidades de una criatura para atribuir estados mentales a otros individuos,

basándose para ello en la observación de su comportamiento en un determinado contexto, con el objetivo de explicar y predecir su conducta (Goldman y Sripada, 2005; Hutto, Herschbach y Southgate, 2011).

En este trabajo, voy a centrarme en una habilidad específica que forma parte de la atribución intencional denominada “reconocimiento de emociones”. Esta capacidad consiste en la detección y atribución de emociones a otros individuos sobre la base de sus expresiones faciales y corporales. Esta capacidad está basada de manera predominante en claves visuales o auditivas, que integran información proveniente del rostro, la posición del cuerpo y la voz. Por ello, en este tipo de habilidades se destaca la relevancia de algún tipo de interacción cara a cara que le permita al intérprete poder extraer ese tipo de información. Específicamente, voy a sostener que la percepción dinámica y contextualizada de las expresiones faciales afina nuestras habilidades de reconocimiento emocional.

De acuerdo con la literatura especializada, el reconocimiento de emociones está definido por dos características particulares: a) el agente interpretado despliega algún patrón de comportamiento expresivo (por ejemplo, una expresión de enojo), y b) el intérprete debe tener algún tipo de acceso observacional y/o perceptual a las expresiones faciales y/o corporales del agente interpretado. En este trabajo, solo me voy a concentrar en aquellos casos de reconocimiento de emociones donde existe una interacción cara a cara entre intérprete e interpretado y donde el interpretado manifiesta un tipo de comportamiento expresivo observable para el intérprete. Estas aclaraciones son importantes debido a que algunos autores han distinguido entre *reconocer (emotion recognition)* y *conocer (emotional knowledge)* una emoción (Castro, Cheng, Halberstadt y Grühn, 2015). Las habilidades para reconocer emociones

requieren detectar las expresiones prototípicas y las claves contextuales relevantes para la emoción en cuestión. Por el contrario, las capacidades de “conocimiento” requieren de habilidades lingüísticas y conceptuales complejas y suelen estar sustentadas en vínculos entre intérprete e interpretado que no involucran necesariamente una interacción cara a cara⁵⁴ (por ejemplo, a partir de testimonios de un tercero). Además, según Grühn, Lumley, Diehl y Labouvie-Vief (2013), las capacidades de conocimiento de las emociones suelen involucrar capacidades complejas, como entender las funciones sociales y morales de las emociones y la monitorización de estados representacionales propios y ajenos. Posiblemente, este tipo de capacidades requieran de un análisis diferente al de las habilidades de reconocimiento, que no voy a discutir en este trabajo⁵⁵.

Distintos enfoques de la atribución intencional han intentado explicar esta capacidad. Podemos distinguir entre ellos dos grandes grupos: a) las *posturas inferencialistas* y b) las *no inferencialistas*. Entre las primeras se ubican diversas versiones de la *teoría de la teoría* y de la *teoría de la simulación*, dos enfoques clásicos que sostienen que hay una brecha entre la observación del comportamiento

⁵⁴ Por ejemplo, cuando un intérprete intenta comprender las emociones del personaje principal de una novela, es probable que tenga que apelar a información guardada en la memoria, construir alguna narrativa acerca del personaje y recurrir a algún tipo de conocimiento conceptual, lingüístico y proposicional para poder comprender las representaciones internas del personaje que lo hacen sentir de tal o cual manera.

⁵⁵ Aunque Castro y colaboradores no dicen nada al respecto, considero que podría establecerse una línea de continuidad gradual entre los casos de reconocimiento de emociones y los casos de conocimiento. Es posible que haya situaciones en nuestras interacciones cotidianas en las que las habilidades de reconocimiento y las de conocimiento estén solapadas, conformando casos de comprensión emocional “híbridos”. Ahora bien, como he señalado, en este trabajo solo voy a concentrarme en los casos de reconocimiento de emociones, dejando de lado los casos de conocimiento y los posibles casos de solapamiento entre los dos tipos de habilidades.

(las expresiones faciales) y el estado mental subyacente del interpretado (la emoción)⁵⁶. De acuerdo con estas posiciones, esta brecha es cubierta por medio de algún proceso de inferencia, ya sea empleando algún tipo de teoría psicológica *folk* (TT) o mediante un proceso de simulación (TS). Dentro del segundo grupo, se encuentran aquellos que han defendido el enfoque de la *percepción directa* en sus diferentes variantes. Los defensores de este tipo de propuesta han puesto de relieve que hay casos de interacción que son difíciles de explicar desde las perspectivas ortodoxas. Estos tienen la particularidad de que no requieren necesariamente de elementos lingüísticos (aunque puede haberlos) y de que descansan, fundamentalmente, en un involucramiento corporal y emocional, tanto del intérprete como del interpretado, donde el comportamiento es “leído” como significativo o expresivo en distintos tipos de prácticas sociales. Por ejemplo, las interacciones primarias entre un bebé y su cuidador, los movimientos de los jugadores de un equipo de fútbol, una coreografía de una pareja de baile, etc. (Pérez, 2013). Brevemente, estos sostienen que nuestro modo primario de interactuar no requiere de una lectura de la mente del modo en que las teorías tradicionales habían pensado, sino que descansa en la percepción directa de, al menos, algunos estados mentales. De acuerdo con estos autores, el intérprete ve la tristeza en el

⁵⁶ La teoría de la teoría (TT), supone que la atribución mental está mediada por una gran cantidad de información y conocimiento conceptual a partir del cual se configura una “teoría” acerca de los estados mentales ajenos. Para este enfoque, la predicción y explicación del comportamiento ajeno está estrechamente vinculada con el empleo, por parte del intérprete, de una serie de leyes que vinculan los estados mentales con el comportamiento (Gopnik y Meltzoff, 1997; Wellman, 1990). Por su parte, la teoría de la simulación (TS) sostiene que la comprensión, explicación y predicción de los estados mentales y el comportamiento ajeno está mediada por un mecanismo de simulación por el cual el intérprete “finge” o “pretende” poseer los estados mentales del interpretado con el objeto de predecir cuál será su acción posterior, empleando su propio aparato cognitivo (Heal, 1996; Goldman, 2005; Gordon, 1986).

rostro, observa la vergüenza en la coloración de sus mejillas, percibe el miedo o la ansiedad en su voz, etc⁵⁷. Básicamente, este enfoque rechaza que las inferencias entre el comportamiento y el estado mental jueguen un rol fundamental en la comprensión de los otros.

Pese a los grandes desacuerdos respecto a cómo explicar en qué consisten las habilidades de reconocimiento y atribución emocional y qué capacidades cognitivas e interactivas involucran, existe una tendencia general entre los investigadores a considerar las expresiones faciales y corporales como una fuente fundamental de información al momento de atribuir un estado emocional. Muchas de las propuestas teóricas contemporáneas se apoyan en los aportes tradicionales de Ekman y cols. (Ekman 1993; 1999), quienes sostienen que algunas emociones, denominadas “básicas”, poseen expresiones faciales y corporales prototípicas que se presentan en todas las culturas. Específicamente, el equipo de Ekman sostiene que todos los miembros de nuestra especie expresan e interpretan las expresiones de las emociones básicas de manera similar, independientemente de su contexto y de las variables socioculturales, debido a que estas tienen un valor comunicativo universal.

En el marco de estos debates, voy a sostener que la dinámica del comportamiento expresivo, entendida como la secuencia temporal de las expresiones faciales y corporales, es un factor esencial al momento de explicar la habi-

⁵⁷ La idea de que el mecanismo de atribución primaria de estados mentales es perceptual ha sido complementada y enriquecida con otras tesis, como la de la interacción cara a cara (Newen, Welpinghus y Juckel, 2015) y la del carácter extendido, en activo y corporizado de la mente. Este conjunto de tesis ha dado lugar a un enfoque más amplio y actual denominado “de segunda persona o interaccionista”. Para más detalles ver Gomila (2002), Scotto (2002), Gallagher (2008) y Clark (2010).

lidad para reconocer emociones⁵⁸. Además, voy a defender que la dinámica de la expresión facial, en conjunto con el registro de cierto tipo de información contextual, colabora para una detección apropiada de las emociones que los otros experimentan. A fin de alcanzar el objetivo propuesto, seguiré los siguientes pasos: en primer lugar, voy a examinar brevemente algunas de las metodologías usuales que los psicólogos cognitivos han empleado para evaluar el reconocimiento de emociones. En segundo lugar, voy a mencionar algunas dificultades que afectan a dichos estudios tradicionales, principalmente basadas en la evaluación del reconocimiento de emociones a partir de expresiones faciales desprovistas de un contexto. Por último, a partir de los aportes de Stout (2010) acerca de la importancia de la dinámica de las expresiones faciales, señalaré que la dinámica, en conjunto con el contexto, permite al intérprete identificar adecuadamente el tipo de estado emocional y, además, posibilita que pueda diferenciarlos de otro tipo de estados psicológicos.

2. El estudio del reconocimiento de emociones

Durante las últimas décadas, la psicología cognitiva, la psicología del desarrollo y las neurociencias han utilizado diferentes metodologías para abordar el estudio de la

⁵⁸ Como he mencionado en la introducción, voy a restringir la dinámica de las expresiones a la secuencia de movimientos faciales y corporales, dejando de lado aquellos casos en los que haya, además, interacciones verbales entre los sujetos u otro tipo de procesos de comprensión emocional involucrados que exceden a la definición de “reconocimiento de emociones” proporcionada más arriba.

habilidad para reconocer emociones⁵⁹. Debido a la preponderancia que posee la información facial en las interacciones sociales entre los individuos, muchas de las pruebas estuvieron enfocadas en evaluar la detección de las emociones por medio de las expresiones del rostro o partes de él (Rhodes, Calder, Johnson y Haxby, 2011).

Los enfoques de corte evolucionista de las emociones han sido pioneros en el estudio científico de esta habilidad (Ekman y Friesen, 1999; Matsumoto *et al.*, 2000). Este tipo de perspectivas sostienen, al menos, dos tesis principales: a) que existe un conjunto de emociones básicas que son de carácter universal, las que, según Ekman, serían alegría, tristeza, asco o disgusto, sorpresa, miedo y enojo; y b) esas emociones básicas son universalmente reconocidas, independientemente de la cultura donde esté inmerso el individuo, debido al valor evolutivo y comunicativo de dichos estados⁶⁰.

Inspirado en los aportes de Darwin (1872), el equipo de Paul Ekman inauguró una corriente metodológica que evalúa el reconocimiento de emociones basado en la

⁵⁹ Luego, las pruebas comenzaron a diversificarse, y se evaluó la capacidad para detectar emociones por medio de los tonos de la voz (Scherer y Scherer, 2011) o a través de videos con personajes que interactúan entre sí, donde se demanda que el sujeto comprenda no solo los estados emocionales, sino también creencias, deseos y otros estados cognitivos (Dziobek *et al.*, 2006).

⁶⁰ Recientemente, Ekman (2003) ha sostenido que podemos identificar las emociones básicas como “programas de afecto”, esto es, una familia de estados relacionados, donde cada miembro de la familia comparte algunas características generales de la emoción en cuestión, pero admite algunas variaciones. Por ejemplo, la emoción de alegría podría subsumir a una familia de estados similares, como la gratificación, el éxtasis, la diversión, etc. De acuerdo con Ekman, cada emoción tiene un “tema” y una “variación”. El tema está compuesto por las características únicas de esa familia, mientras que las variaciones son productos de las diferencias individuales y las del contexto específico donde ocurre esa emoción. Los temas son producto de la evolución, y las variaciones, producto del aprendizaje. Los programas de afecto son complejos, evolutivamente determinados, responden a una variedad de condiciones ambientales y pueden incluir expresiones faciales, cambios fisiológicos, evaluaciones y disposiciones conductuales.

detección de patrones de expresiones del rostro y, posteriormente, de algunas partes del cuerpo. Brevemente, han defendido la idea de que la detección de emociones es un proceso rápido y automático asociado a claves prototípicas de cada emoción básica. Para ello, desarrollaron el método FACS (*Facial Action Coding System*, o, en español, sistema de codificación facial), donde se establecieron las expresiones prototípicas de cada emoción básica y se desarrollaron métodos para la evaluación del reconocimiento de dichas expresiones por parte de otros individuos. Este sistema está basado en la anatomía del rostro y registra cada movimiento de los músculos faciales (al que denominan “unidad de acción”) subyacente a la expresión de las emociones básicas. El sistema distingue 44 unidades de acción, que consisten en la combinación de diferentes movimientos de los músculos faciales, posiciones de la cabeza, dirección de la mirada, etc., codificando cada movimiento y luego asignándole a cada combinación un número que la identifica. Específicamente, el sistema de codificación facial tiene como finalidad poder medir objetivamente aquellas subexpresiones del rostro y lograr una interpretación lo más exacta posible de la expresión facial en su totalidad. El FACS ha sido uno de los sistemas para evaluar reconocimiento de emociones más utilizado, ya que constituye una guía que permite recrear expresiones emocionales prototípicas y estandarizadas a partir de indicarles a los sujetos los movimientos propios de cada unidad de acción inherente a la expresión de, al menos, las emociones básicas. Esto es, se les pide a sujetos voluntarios que realicen los movimientos faciales propios de cada emoción básica y luego se les toma una serie de fotografías. Estas imágenes son las que posteriormente se utilizan como material para evaluar el reconocimiento de emociones.

Específicamente, utilizando este sistema, Ekman tomó 110 fotografías de expresiones faciales, teniendo en cuenta las unidades de acción mencionadas anteriormente: las fotografías de afectos faciales (POFA). Estas han sido utilizadas en diversos campos de investigación, especialmente en estudios transculturales, con el objetivo de probar la universalidad de las emociones a partir de su reconocimiento. Como he mencionado anteriormente, Ekman sostiene que las emociones básicas tienen valor adaptativo, razón por la cual todos los miembros de la especie expresan y reconocen las expresiones de las emociones básicas de manera similar, independientemente de su contexto y de las variables socioculturales. Dado su carácter fundamental para la supervivencia, la expresión de las emociones básicas es rápidamente captada o registrada por otros individuos. Según Ekman, el asco se caracteriza por la nariz arrugada y el labio superior levantado; el miedo, por las cejas levantadas y juntas, los párpados superiores levantados y los inferiores tensos, y la boca ligeramente estirada; la sorpresa, en cambio, se caracteriza por las cejas elevadas y los ojos y boca bien abiertos. Según su perspectiva, este tipo de expresiones da indicios a nuestros semejantes de lo que le ocurre psicológicamente a quien experimenta la emoción. Además, afirma que la observación de las expresiones faciales podría proporcionar algún tipo de información acerca de los antecedentes de la emoción o de sus causas, como también de algunos comportamientos posteriores de quien las posee. Por ejemplo, cuando vemos una expresión de asco, sabemos que la persona está respondiendo a “algo”, por ejemplo, un olor o un gusto desagradable, que es probable que haga sonidos de desagrado antes que de agrado y que intente evadir la fuente de aquello que le resulta desgradable (Ekman, 1999,

p. 2). Así pues, este tipo de información tendría un valor importante en la regulación y coordinación de las relaciones interpersonales.

Si bien este tipo de enfoques ha tenido una gran recepción entre los psicólogos y filósofos de la mente, las críticas no tardaron en llegar. Gran parte de ellas están relacionadas con un debate clásico respecto al peso relativo que las expresiones faciales tienen al momento de atribuir estados emocionales frente a otras variables, por ejemplo, contextuales (Carroll y Russell, 1996). Los enfoques evolucionistas del reconocimiento de emociones han sostenido la tesis de la “dominancia facial”, que afirma que, cuando un intérprete ve una expresión facial de una emoción básica, la información del rostro se vuelve tan saliente y preponderante que el intérprete desestima otro tipo de información, como, por ejemplo, la situación en la que se encuentra el individuo que expresa la emoción o las expectativas previas del intérprete (Buck, 1984; Ekman, 1972; Fridlund, Ekman y Oster, 1987; Izard, 1971). Otros, en cambio, defienden la tesis de la “dominancia limitada”, que sostiene que las expresiones faciales no son la única variable que los individuos tienen en cuenta al momento de atribuir la emoción. Según estos autores, la información situacional juega un rol importante en la atribución (Fridlund, 1994; Ortony y Turner, 1990; Russell, 1994; Woodworth, 1938).

Estudios recientes favorecen, con distintos matices, la línea de la dominancia limitada de las expresiones faciales, y han señalado algunos límites teóricos y metodológicos de aquellas corrientes que ponderan en exceso el papel de las expresiones faciales. De acuerdo con sus detractores, una de las desventajas de este tipo de metodologías es que la detección de las emociones ha sido investigada usando fotos estáticas de expresiones faciales que maximizan

la distinción entre categorías y generan resultados experimentales forzados y artificiales. En nuestra vida diaria, las expresiones faciales están fuertemente influenciadas por la cultura y no son tan nítidas, rígidas y prototípicas. En segundo lugar, se ha señalado que la detección de emociones está siendo evaluada de manera descontextualizada. Habitualmente, los rostros no son percibidos de manera aislada, sino que aparecen en un contexto particular, esto es, en un ambiente específico, en conjunto con otras personas y, además, de manera concomitante, con otro tipo de información socialmente significativa. Aún más, existe un conjunto de información de carácter social, adquirida por medio del aprendizaje y de distintos tipos de procesos cognitivos (por ejemplo, prejuicios raciales), que juega un rol en la percepción de los estímulos emocionalmente importantes, por ejemplo, dirigir nuestra atención (Feldman Barrett *et al.*, 2011). En conclusión, la percepción de las expresiones faciales está influenciada por diferentes tipos de variables, más allá de la información estrictamente facial.

Ahora bien, veamos con algo más de detalle qué significa que las variables contextuales inciden en el reconocimiento de emociones a partir de las expresiones faciales. Habitualmente, en la literatura, la expresión “variables contextuales” puede aludir a múltiples ideas. Frecuentemente, alude a variables situacionales relacionadas con la escena particular donde se desarrolla la atribución psicológica y la situación en la que se encuentra el individuo que experimenta la emoción; esto es, en qué situación la persona está desplegando tal o cual patrón expresivo, por ejemplo, en un partido de fútbol, en un cumpleaños, en el trabajo, etc. A veces, las mismas expresiones faciales en contextos distintos son reconocidas como alusivas a diferentes emociones (Feldman Barrett *et al.*, 2011). En este

artículo, cuando me refiera a “variables contextuales” voy a hacerlo en este sentido, que es el cual se ha utilizado habitualmente.

En otras ocasiones, “contextual” refiere a las características del rostro y de la posición del cuerpo que no tienen que ver estrictamente con el modo prototípico en el que se manifiesta este conjunto de emociones básicas, por ejemplo, la dirección de la mirada, la postura corporal y la prosodia afectiva. Algunos estudios muestran que la dirección de la mirada, por ejemplo, suele estar relacionada con la intensidad de la emoción que los sujetos atribuyen, mientras más directa es la mirada del individuo (en oposición a la mirada apartada o distraída), más intensa es la emoción que se atribuye. Por otro lado, los sujetos ven disminuido su éxito en tareas de reconocimiento cuando la mirada del interpretado se encuentra apartada (Binde-mann *et al.*, 2008).

En un tercer sentido, “contextual” refiere a variables relacionadas con el intérprete: descripciones verbales previas que tenga del contexto, sesgos de distintos tipos (por ejemplo, raciales o de género), ciertos tipos de aprendizaje social, tales como: “Las personas cuando están tristes hacen tales o cuales cosas”, entre otras cosas (Wieser y Brosch, 2012).

Esta caracterización acerca de los diferentes tipos de variables contextuales no tiene la intención de ser exhaustiva, sino de mostrar los motivos más relevantes por los cuales se ha justificado la tesis de la dominancia limitada de las expresiones faciales. Teniendo en cuenta estas consideraciones respecto a la complejidad de las emociones y su reconocimiento, en el siguiente apartado voy a argumentar a favor de que la dinámica de las expresiones faciales y el contexto en el que se expresa la emoción son dos elementos fundamentales para dar cuenta

de la habilidad para reconocer emociones. Por “dinámica de las expresiones faciales” voy a entender a la secuencia temporal de movimientos faciales y corporales que se dan durante la expresión de ciertos estados psicológicos. En la mayoría de nuestras interacciones cotidianas, los sujetos con los que interactuamos despliegan movimientos faciales y corporales continuos que nos permiten extraer distinto tipo de datos socialmente relevantes. Por ejemplo, nos proporcionan información acerca de su foco de atención (Emery, 2000; Nummenmaa y Calder, 2009) y, según Zlateanu, Krumhuber y Richardson (2018), la dinámica de los movimientos faciales permite mejorar nuestra habilidad para reconocer expresiones de emociones improvisadas versus emociones genuinas⁶¹.

Pese a que distintos aportes empíricos han reconocido la importancia de la dinámica de las expresiones en el proceso para reconocer emociones, considero que no se ha examinado teóricamente con suficiente atención cuál es el aporte que efectivamente tiene la dinámica de las expresiones para este tipo de habilidades de atribución intencional.

⁶¹ En este estudio, los autores evaluaron el reconocimiento de emociones en tres situaciones experimentales: a) en la que las personas expresaban sorpresa genuina; b) en la que se le pedía a los sujetos que improvisaran o fingieran una expresión de sorpresa; y c) en la que se les pedía que fingieran sorpresa pero luego de haber pasado por la primera situación descripta, es decir, de haber experimentado sorpresa genuina. En cualquiera de los tres casos, la dinámica de las expresiones faciales les permitió a los sujetos experimentales reconocer, con mayor éxito, si las expresiones emocionales de las personas que observaban eran improvisadas, falsas o genuinas.

3. La dinámica de las expresiones faciales y el reconocimiento de emociones

En este último apartado, voy a centrarme en dar razones a favor de la importancia de la dinámica de las expresiones faciales para el reconocimiento de emociones. Para ello, voy a retomar algunos aportes filosóficos de Stout (2010) acerca de la detección de estados emocionales que reparan en la distinción proceso-producto, y destacan la importancia de la dinámica temporal de las expresiones. Específicamente, Stout argumenta a favor de la relevancia de la dinámica de las expresiones en el marco de una defensa del enfoque de la percepción directa de los estados mentales y, en este caso, emocionales. Si bien en este artículo no me ocupo del enfoque de la percepción directa de las emociones, ni de la defensa particular que el autor hace de ella, solo voy a hacer referencia a la percepción directa con el objetivo de contextualizar y esclarecer la posición de Stout.

En el artículo “*Seeing the Anger in Someone’s Face*”, Stout defiende la tesis de que podemos percibir directamente, al menos, algunos estados mentales (Gallagher, 2008; Gallagher y Varga, 2014; Krueger, 2012; McNeill, 2012; Pacherie, 2005; Smith, 2015; 2016 y Zahavi, 2011). Específicamente, estos enfoques sostienen que, en algunas ocasiones, podemos tener un acceso perceptual a los estados mentales sin necesidad de emplear inferencias u otros procesos cognitivos de orden superior. En efecto, como ya señalamos antes, los defensores de la percepción directa rechazan la idea de que todos los casos de atribución mental sean casos de transiciones inferenciales entre el comportamiento y el estado mental o transiciones entre contenidos proposicionales. En algunas ocasiones, usamos intuitivamente expresiones que parecen dar cuenta de este

fenómeno: “Veo que Juan está enojado”, “Siento la ansiedad en su voz”, “Veo la vergüenza en la coloración de sus mejillas”, “¿No ves acaso que está feliz?”, etc.

Para entender la estrategia que Stout emplea para defender la percepción directa, es necesario mostrar que las tesis que este enfoque propone pueden abordarse de diferentes modos. Por un lado, en un *sentido psicológico*: como una tesis acerca de los mecanismos que utilizamos para atribuir estados mentales, en este caso perceptuales (en oposición a inferenciales, como en el caso de la teoría de la teoría o de la teoría de la simulación, donde el intérprete realiza una serie de inferencias entre el comportamiento, el ambiente y el estado mental)⁶². Por el otro, en un *sentido epistemológico*, que alude a si efectivamente podemos conocer las mentes de los otros por vías perceptuales. Por último, en un *sentido fenomenológico*, que refiere a cómo las mentes son percibidas conscientemente, su carácter experiencial. Específicamente, Stout se va a centrar en el aspecto epistemológico de la percepción directa, aunque también va a hacer algunas alusiones a la percepción directa en sentido psicológico, que es el aspecto que me interesa en este artículo⁶³.

⁶² En tiempos recientes, algunos filósofos (Hershbach, 2015; Michael y De Bruin, 2015) han caracterizado a la percepción como un proceso inferencial. Esta idea ha generado múltiples debates entre aquellos que sostienen a la percepción directa en un sentido psicológico. Los que defienden esta idea, sostienen nociones de inferencia deflacionada o minimalista, de carácter sub-personal, implícita y no proposicional diferente a la noción de inferencia en los debates clásicos sobre la atribución intencional.

⁶³ Las razones por las que es importante distinguir los distintos modos de entender la percepción directa son variadas, por ejemplo, algunos autores han puesto de relieve que, aunque la fenomenología del intérprete podría ser la de una detección inmediata, directa y no inferencial, esto no excluye que en el curso de la atribución estén involucrados procesamientos de tipo sub-personal o algún otro tipo de proceso explicado en términos computacionales (Spaulding, 2015; Herschbach, 2015).

Conjuntamente con las discusiones que se han dado en el ámbito de la filosofía de la mente acerca de si podemos acceder perceptual o inferencialmente a los estados emocionales de otros individuos, ha habido un gran cuerpo de evidencia empírica, proveniente de la psicología cognitiva y de las neurociencias, que ha sido integrada a estos debates para fortalecer o rechazar argumentos y para mostrar algún punto importante acerca de las diferentes posiciones. Los abordajes evolucionistas y neuroculturales de las emociones, bien establecidos dentro de la psicología, han sido retomados por los defensores de la percepción directa⁶⁴ porque ponen de relieve dos ideas que son interesantes para este enfoque. Por un lado, la tesis de que existe una conexión estrecha entre la emoción y sus expresiones faciales, razón por la cual puede ser más plausible que pudiéramos “observarlas” o tener algún tipo de acceso perceptual⁶⁵. Por el otro, la idea de que el reconocimiento de emociones es una habilidad automática, relativamente simple e independiente de la cultura donde se inserta el individuo. Esto último armoniza bien con el enfoque de la percepción directa porque uno de sus focos de atención radica en explicar interacciones primitivas, tempranas y propias del desarrollo ontogenético. En este contexto, una gran cantidad de filósofos y psicólogos que realizan estudios teóricos sobre la habilidad para reconocer emociones han retomado estos experimentos y elaborado distintos tipos de propuestas.

⁶⁴ También, este tipo de pruebas ha sido utilizado por otro tipo de enfoques de la atribución intencional, como la teoría de la simulación aplicada al reconocimiento de emociones (Goldman, 2006).

⁶⁵ Existen dos debates en torno a este punto que, si bien están estrechamente emparentados, es necesario diferenciar: a) la discusión acerca de cuán constitutivos y universales son los patrones de expresiones faciales a la emoción experimentada por el sujeto (Jack *et al.*, 2012); y b) cuán relevantes son los patrones de expresión facial y/o corporal para una adecuada detección de la emoción por parte del intérprete (Russell, 199; Nelson y Russell, 2013).

No obstante, la idea del reconocimiento de emociones como una habilidad universal y basada exclusivamente en la percepción de expresiones faciales ha sido desafiada en varias oportunidades. Desde el punto de vista de Stout, aquellas explicaciones acerca del reconocimiento de emociones que se basan en las pruebas tradicionales provenientes de la línea de investigación de Ekman se ven enfrentadas a algunos problemas, pero por razones ligeramente diferentes a las que señalé anteriormente (relativas a la evaluación de la habilidad desprovista de variables contextuales). Retomando la idea del apartado anterior, la metodología habitual consiste en mostrarles a los sujetos fotos y pedirles que identifiquen de qué estado emocional se trata. Así, lo que se espera es que las respuestas sean similares independientemente de las diferencias interindividuales y culturales. Esto ha llevado a algunos filósofos a pensar que las emociones básicas tienen una especie de *apariencia característica* (Smith, 2015) o *componente característico* (Green, 2010) que hace que cada una de las emociones básicas se “vea” de una cierta manera que es sensible a nuestras capacidades cognoscitivas⁶⁶.

Una de las críticas principales de Stout refiere a una preocupación epistemológica acerca de si podemos distinguir de manera confiable las emociones a partir de la observación de patrones de expresión facial en una fotografía. De acuerdo con Stout, no podemos. Según él, los autores que basan sus explicaciones en este tipo de estudios no distinguen entre asociar un patrón de expresión facial con una emoción y reconocer realmente cuál es el estado emocional. La razón es que los observadores logran

⁶⁶ Según Green, podemos identificar un conjunto de expresiones que funcionan como *componentes característicos*, es decir, un aspecto de ese estado emocional que habilita al intérprete a atribuir apropiadamente la emoción en cuestión en un contexto particular. Para más detalles, ver Green (2010; 2007).

reconocer el *tipo* emocional en un patrón facial (por ejemplo, enojo, alegría, tristeza, sorpresa) independientemente de si la imagen corresponde a una emoción real o a una falsa. De hecho, las pruebas están realizadas con actores voluntarios que representan el estado emocional según las indicaciones de los investigadores, y no con personas que genuinamente experimentan esos estados. Sin embargo, la pregunta en cuestión es si uno puede realmente ver el estado emocional real de alguien en una expresión facial estática. Para Stout, estos experimentos solo demuestran que somos buenos para asociar una expresión facial con una emoción, pero no evidencian patrones confiables y flexibles de reconocimiento.

Para dar sentido a su idea, Stout hace una analogía del caso del reconocimiento de emociones con la firma de alguien en un papel (*signature*). Cuando alguien firma en una hoja de papel, realiza un escrito que representa su nombre. Sin embargo, esta no sería una firma propiamente dicha si los patrones de la escritura fueran, por ejemplo, accidentales o falsificados. Así pues, solo con mirar una firma en un papel, uno no puede ver si realmente es una firma auténtica, solo puede inferir que es tal a partir de una o varias de sus características. Según el autor, algo similar sucede cuando nos enfrentamos a una expresión de enojo. Si solo vemos un rostro con una expresión de enojo, no podemos distinguir si es una emoción real o una falsa. De hecho, un buen actor puede engañarnos simulando una buena expresión de enojo, sin estar experimentando en absoluto una emoción. En resumen, la analogía es simple, de la misma manera que no podemos saber con solo ver la firma de alguien en un papel si es una firma genuina o falsa, no podemos conocer fiablemente el enojo de alguien con solo ver la foto, debido a que no podemos distinguir si esta es falsa o genuina.

Para Stout, lo que debemos agregar para poder dar cuenta del reconocimiento de emociones de manera fiable es que el intérprete pueda captar el *proceso dinámico* mediante el cual se despliega la emoción. No basta con captar *productos* parciales del proceso, como las imágenes estáticas capturadas por una fotografía, porque estas solo muestran asociaciones simples entre expresiones y estados mentales. Por el contrario, lo que necesitamos es observar el despliegue secuencial de las expresiones corporales y faciales para acceder de manera confiable a ese estado mental. Como he señalado anteriormente, captar la dinámica temporal es parte de captar el proceso mediante el cual se expresa la emoción. Básicamente, se trata del registro, momento a momento, de los cambios relevantes en los patrones de expresión facial y corporal de un individuo (Sato, 2004).

Además del problema de distinguir emociones de manera confiable, existe otra dificultad teórica relacionada con la evaluación del reconocimiento mediante imágenes estáticas que Stout menciona, brevemente y sin demasiada atención, pero que también pone de relieve la importancia de la dinámica del comportamiento expresivo. El autor señala que reparar en el proceso dinámico en el que se está desplegando un estado emocional nos permite distinguir finamente ante qué *tipo* de estado mental estamos (lo que técnicamente se ha denominado en filosofía de la mente como la “actitud psicológica”). Una de las razones que Stout esgrime a favor de considerar el proceso es que hay ciertos elementos expresivos que, en principio, parecerían característicos de ciertas emociones (por ejemplo, el llanto en la tristeza, la sonrisa en la alegría o el ceño fruncido en el enojo), pero que se presentan en distintos tipos de estados, incluso no emocionales. Por ello, es preciso considerar en qué contexto y en qué proceso se

despliegan. Por ejemplo, estallar en lágrimas de frustración es un proceso distinto del sollozo propio de la tristeza, el rubor por vergüenza es un proceso distinto del enrojecimiento de las mejillas debido a los efectos del alcohol en los vasos sanguíneos, o retorcer la cara (*twitching the face*) de pena es distinto de retorcer la cara por desesperación o por dolor (Stout, 2010, p. 39). En estos casos, la dinámica de las expresiones es una buena herramienta para poder distinguir con mayor fineza ante qué *tipo* de estado mental estamos, porque nos permite captar los pasos previos y los posteriores y, además, observar cuál es el curso específico que toma el comportamiento expresivo vinculado con los distintos estados mentales.

Esta indeterminación o ambigüedad de ciertos patrones expresivos también ocurre, por ejemplo, con los rostros sonrientes. Orlowska, Krumhuber, Rychlowska y Szarota (2018) han señalado que, a pesar de la asociación entre sonrisas y sentimientos positivos como la alegría, también estos patrones se presentan de modo similar en situaciones donde se quiere expresar cooperación y confianza. Incluso, puede mostrarse también durante experiencias desagradables, como un modo de ocultar sentimientos negativos o mostrar menor estatus social ante una figura de autoridad. Por lo tanto, la sonrisa se utiliza en una amplia variedad de situaciones, según el contexto y las normas sociales aprendidas a través de la socialización y la experiencia.

Asimismo, fuera del contexto de laboratorio, las expresiones emocionales no se presentan de manera prototípica, como inicialmente habían pensado los investigadores. En su mayoría, las expresiones faciales en contextos cotidianos de interacción están constreñidas por diversos factores sociales y culturales, razón por la cual se ven mezcladas con gestos idiosincrásicos de cada individuo y otros

propios de la cultura en la que se encuentra inmerso. Además, las expresiones emocionales van cambiando a una gran velocidad en un período corto de tiempo, ajustándose a los cambios del contexto. Una vez más, reparar en el proceso dinámico podría ayudar a desambiguar estos casos, porque permite tener una visión más integrada respecto a cómo tiene lugar la emoción, ante qué está reaccionando ese individuo y cómo va regulando su estado emocional conforme pasa un período de cierto tiempo.

Debido a que las expresiones faciales pueden estar vinculadas con diferentes estados emocionales, la dinámica en conjunto con el contexto proporciona una buena herramienta para dirimir el peso de la información de cada elemento. De acuerdo con Stout, captar el proceso en el que se despliega una emoción implica poder registrar perceptualmente las distintos tipos de expresiones en una secuencia temporal. Pero, además, Stout va un paso más adelante diciendo que esa percepción de las expresiones faciales no solo registra el carácter dinámico, sino que además lo hace de manera activa e interactiva. En consonancia con los aportes de Gibson (1966), Stout recalca que la percepción no es un proceso pasivo, sino que requiere del sujeto que se involucre activamente en el entorno.

Así pues, cuando alguien está enojado podría señalizar su emoción al otro de diferentes maneras: con su expresión facial, pero también elevando la voz y gritando, e imponiendo cierta presencia física. Esto, de alguna manera, interpela y exige una respuesta por parte de su interlocutor. Esa réplica podría ser una disculpa y, de este modo, el enojo podría, eventualmente, disminuir, o bien podría ser otra respuesta de enojo, por la cual la emoción podría intensificarse aún más. Por ello, para Stout, la percepción de las expresiones emocionales implica necesariamente involucrarse en prácticas sociales, que son las

que permiten captar finamente no solo el tipo de estado, sino también aspectos como la intensidad de la emoción y la valencia.

Ahora bien, hasta aquí he mostrado la importancia que podría tener para un intérprete reparar en la dinámica de las expresiones faciales para identificar con mayor fineza el tipo de estado psicológico, en este caso emocional, de otro individuo. Además, he señalado que la detección de esa dinámica se realiza en muchas ocasiones mediante la percepción activa e interactiva del intérprete, que también le permita captar rasgos como la valencia y la intensidad de la emoción. Asimismo, en nuestras interacciones ordinarias, el registro de esa dinámica del comportamiento expresivo se da en contextos y situaciones particulares que nos sugieren diferentes modos de entender la secuencia de las expresiones como perteneciente a un tipo de estado emocional o bien a otro tipo de proceso psicológico.

En resumen, el reconocimiento de emociones está integrado a una variedad de procesos dinámicos de interacción y otros procesos cognitivos, de manera que nunca ocurre de forma aislada y con propósitos epistémicos desinteresados. Finalmente, considero que este conjunto de factores nos permite tener una imagen más integrada teóricamente y empíricamente mejor informada de un fenómeno socialmente muy complejo que posee una importancia social crítica.

4. Conclusiones

En este trabajo, he ofrecido algunas razones para considerar la dinámica de las expresiones faciales como un factor importante al momento de reconocer y atribuir estados emocionales. Para lograr dicho propósito, retomé los

estudios clásicos basados en la detección de emociones por medio de fotografías con expresiones estáticas, provistas por los enfoques evolutivos y neuroculturales de las emociones. Este tipo de estudios ha tenido una influencia extraordinaria en la psicología de la emoción y ha sido fuente de innumerables trabajos teóricos y metodológicos. Ahora bien, en tiempos recientes, han surgido críticas de diversa índole que han permitido examinar y refinar estos protocolos de reconocimiento de emociones. Gran parte de las críticas estuvieron orientadas a que este reconocimiento se evaluaba con expresiones estáticas que estaban desprovistas de un contexto.

En el segundo apartado de este trabajo, señalé posibles acepciones de “variables contextuales” en la literatura y he sostenido que se puede distinguir, al menos, entre: a) variables situacionales, relacionadas con la escena particular donde se desarrolla la atribución; b) aquellas asociadas al agente que experimenta la emoción, como la dirección de la mirada, la postura corporal, la prosodia afectiva entre otras; c) variables referidas al intérprete, como, por ejemplo, información previa emocionalmente significativa y sesgos de razonamiento (v. g. de tipo racial o de género), etc. El objetivo fue mostrar que existe una importancia moderada de las expresiones faciales en algunos contextos de atribución de emociones, y advertir la relevancia de los factores contextuales que ajustan y enriquecen la atribución.

Luego, en el tercer apartado, sostuve que la dinámica de las expresiones faciales es un factor fundamental para comprender cómo tiene lugar el reconocimiento de emociones. En particular, esa dinámica del comportamiento expresivo es la que nos permite advertir el proceso en el que están inmersos esos patrones de expresión facial. Además el despliegue secuencial de las expresiones faciales

tiene lugar en un contexto situacional en que también colabora en la detección del tipo emocional o la actitud psicológica.

En este contexto, la importancia de la dinámica de las expresiones faciales, sumada a un enfoque contextual e interactivo, es fundamental para explicar nuestras habilidades de atribución emocional. Considero que estas reflexiones debieran servirnos de base para crear protocolos experimentales más ecológicos, que no solo consideren la dinámica de las expresiones faciales, sino también posibles modos interactivos entre intérprete e interpretado. Estudios de este tipo proporcionarían una versión teóricamente más rica y flexible del fenómeno de reconocimiento de emociones.

Referencias bibliográficas

- Andrews, K. (2012). *Do Apes Read Minds? Toward a New Folk Psychology*. Cambridge Massachusetts: MIT Press.
- Bindemann, M., Burto, A. M. y Langton, S. R. H. (2008). "How do eye gaze and facial expression interact?" *Visual Cognition*, 16(6), pp. 708-733. DOI: 10.1080/13506280701269318
- Bohl, V. y Gangopadhyay, N. (2013). "Theory of mind and the unobservability of other minds". *Philosophical Explorations*, 17(2), 203-222. DOI: 10.1080/13869795.2013.821515.
- Buck, R. (1984). *The Communication of Emotion*. Nueva York: Guilford Press.

- Castro, V., Cheng, Y., Halberstadt, A. y Grühn, D. (2015). "EUREKA! A Conceptual Model of Emotion Understanding". *Emotion Review*, 8(3), pp. 258-268. DOI: 10.1177/1754073915580601.
- Carroll, J. M. y Russell, J. A. (1996). "Do facial expressions signal specific emotions? Judging emotion from the face in context". *Journal of Personality and Social Psychology*, 70(2), pp. 205-218. DOI: 10.1037//0022-3514.70.2.205. ·
- Clark, A. (2010). *Being There: Putting Brain, Body, and World Together Again*. Londres: The MIT Press.
- Darwin, C. (1872). *The Expression of the Emotions in Man and Animals*. Londres: John Murray.
- Dziobek, I., Fleck, S., Kalbe, E., Rogers, K., Hassenstab, J., Brand y Convit, A. (2006). "Introducing MASC: a movie for the assessment of social cognition". *Journal of Autism and Developmental Disorders*, 36(5), pp. 623-636. DOI: 10.1007/s10803-006-0107-0.
- Ekman, P. (1972). "Universals and Cultural Differences in Facial Expressions of Emotion". En J. K. Cole (Ed.), *Nebraska Symposium on Motivation* (pp. 207-283). Lincoln: University of Nebraska Press.
- Ekman, P. (1993). "Facial expressions and emotion". *American Psychological Association*, 48(4), pp. 384-392. DOI: 10.1037/0003-066X.48.4.384
- Ekman, P. (1999). "Facial Expressions". En Dalgleish, T. y Power, M. J. (Eds.), *The Handbook of Cognition and Emotion* (pp. 301-320). Nueva York: John Wiley y Sons Ltd.
- Ekman, P. (2003). *Emotions Revealed: Recognizing Faces and Feelings to Improve Communication and Emotional Life*. Nueva York: Times Books.

- Emery, N. J. (2000). "The eyes have it: the neuroethology, function and evolution of social gaze". *Neuroscience y Biobehavioral Reviews*, 24(6), pp. 581-604.
- Feldman Barrett, L., Mesquita, B. y Gendron, M. (2011). "Context in emotion perception". *Current Directions in Psychological Science*, 20(5), pp. 286-290. DOI: 10.1177/0963721411422522
- Fridlund, A. J. (1994). *Human Facial Expression: An Evolutionary View*. San Diego, California: Academic Press.
- Fridlund, A. J., Ekman, P. y Oster, H. (1987). "Facial Expressions of Emotion: Review of the Literature", 1970-1983. En A. W. Siegman y S. Feldstein (Eds.), *Nonverbal Behavior and Communication* (pp. 143-224). Nueva Jersey: Lawrence Erlbaum Associates.
- Gallagher, S. (2008). "Direct perception in an intersubjective context". *Consciousness and Cognition*, 17(2), pp. 535-543. DOI: 10.1016/j.concog.2008.03.003.
- Gallagher, S y Varga, S. (2014). "Social Constraints on the Direct Perception of Emotions and Intentions". *Topoi*, 33(1), pp. 185-199. DOI: 10.1007/s11245-013-9203-x.
- Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin.
- Goldie, P. (2002). "How we think of others' emotions". *Mind y Language*, 14(4), pp. 394-423. DOI: 10.1111/1468-0017.00118.
- Goldman, A. y Sripada, C. (2005). "Simulationist models of face-based emotion recognition". *Cognition*, 94(3), 193-213. DOI:10.1016/j.cognition.2004.01.005.
- Gomila, A. (2002). "La perspectiva de segunda persona de la atribución mental". *Azafea*, 4, pp. 123-138.
- Gopnik, A. y Meltzoff, A. N. (1997). *Words, Thoughts, and Theories*. Cambridge: MIT Press.

- Gordon, R. (1986). "Folk psychology as simulation". *Mind and Language*, 1(2), pp. 158-171. DOI: 10.1111/j.1468-0017.1986.tb00324.x.
- Green, M. (2007). *Self-Expression*. Oxford: Clarendon Press.
- Green, M. (2010). "Perceiving emotions". *Proceedings of the Aristotelian Society Supplementary*, 84 (1), 45- 61. DOI: 10.1111/j.1467-8349.2010.00185.
- Grühn, D., Lumley, M., Diehl, M. y Labouvie-Vief, G. (2013). "Time-based indicators of emotional complexity: Interrelations and correlate". *Emotion*, 13(2), pp. 226-237. DOI: 10.1037/a0030363.
- Halberstadt, A. G., Denham, S. A. y Dunsmore, J. C. (2001). "Affective social competence". *Social Development*, 10(1), pp. 79-119. DOI: 10.1111/14679507.00150.
- Herschbach, M. (2015). "Direct social perception and dual process theories of mindreading". *Consciousness and Cognition*, 36, pp. 483-497. DOI: 10.1016/j.concog.2015.08.005.
- Michael, J. y De Bruin, M. (2015). "How Direct is Social Perception?". *Consciousness and Cognition*, 36, pp. 373-375. DOI: 10.1016/j.concog.2015.08.005.
- Heal, J. (1996). "Simulation, theory and content". En P. Carruthers y P. Smith (Eds.), *Theories of theories of mind* (pp. 75-89). Cambridge: Cambridge University Press.
- Hurley, S. L. (2008). "The shared circuits model (SCM): How control, mirroring, and simulation can enable imitation, deliberation, and mindreading". *Behavioral and Brain Sciences*, 31, pp. 1-58. DOI: 10.1017/S0140525X07003123.
- Hutto, D. D., Herschbach, M. y Southgate, V. (2011). "Mindreading and alternatives". *Review of Philosophy and Psychology*, 2(3), pp. 375-395. DOI: 10.1007/s13164-011-0073-0.

- Izard, C. E. (1971). *The face of emotion*. Nueva York: Appleton-Century Crofts.
- Jack, R. E., Garrod, O. G., Yu, H., Caldara, R., y Schyns, P. G. (2012). "Facial expressions of emotion are not culturally universal". *Proceedings of the National Academy of Sciences*, 10(19), 7241-7244. DOI: 10.1073/pnas.1200155109.
- Krueger, J. (2012). "Seeing Mind in Action". *Phenomenology and the Cognitive Sciences*, 11(2), pp. 149-173. DOI:10.1007/s11097-011-9226-y.
- Matsumoto, D., LeRoux, J., Wilson-Cohn, C., Raroque, J., Kooken, K., Ekman, P. y Goh, A. (2000). "A new test to measure emotion recognition ability: Matsumoto and Ekman's Japanese and Caucasian Brief Affect Recognition Test (JACBART)". *Journal of Nonverbal Behavior*, 24(3), pp. 179-209. DOI:10.1023/a:1006668120583.
- McNeill, W. E. S. (2012). "On seeing that someone is angry". *European Journal of Philosophy*, 20(4), pp. 575-597. DOI: 10.1111/j.1468-0378.2010.00421.x.
- Nelson, N. L., y Russell, J. (2013). "Universality Revisited". *Emotion Review*, 5(1), pp. 8-15. DOI: 10.1177/1754073912457227.
- Newen, A., Welplinghus, A., Juckel, G. (2015). "Emotion Recognition as Pattern Recognition: The Relevance of Perception". *Mind y Lenguaje*, 30(2), pp. 187-208. DOI: 10.1111/mila.12077.
- Nummenmaa, L., y Calder, A. J. (2009). "Neural mechanisms of social attention". *Trends on Cognitive Science*, 13(3), pp. 135-143. DOI: 10.1016/j.tics.2008.12.006.
- Orlowska, A. B., Krumhuber, E. G., Rychlowska, M. y Szarota, P. (2018). "Dynamics Matter: Recognition of Reward, Affiliative, and Dominance Smiles From Dynamic vs. Static Displays". *Frontiers in Psychology*, 9(938), pp. 1-8. DOI: 10.3389/fpsyg.2018.00938.

- Ortony, A. y Turner, T. J. (1990). "What's basic about basic emotions?". *Psychological Review*, 97(3), pp. 315-331. DOI: 10.1037/0033-295X.97.3.315.
- Pacherie, E. (2005). "Perceiving intentions". En J. Sàágua (Ed.), *A Explicação da Interpretação Humana* (pp. 401-414). Lisboa: Edições Colibri.
- Pérez, D. (2013). *Sentir, desear, creer. Una aproximación filosófica a los conceptos psicológicos*. Buenos Aires: Prometeo.
- Rhodes, G., Calder, A. Johnson, M. y Haxby, J. V. (2011). *Oxford Handbook of Face Perception*. Oxford: Oxford University Press.
- Russell, J. A. (1994). "Is there universal recognition of emotion from facial expression? A review of the cross-cultural studies". *Psychological Bulletin*, 115(1), pp. 102-141. DOI: 10.1037/0033-2909.115.1.102.
- Russel, J. A (1997). "Reading emotions from and into faces: Resurrecting a dimensional-contextual perspective". En J. Russell and J. M. Fernández-Dols (Eds.), *The Psychology of Facial Expression* (pp. 295-320). Cambridge: Cambridge University Press.
- Sato, W. y Yoshikawa, S. (2004). "The dynamic aspects of emotional facial expressions". *Cognition and Emotion*, 18(5), pp. 701-710. DOI: 10.1080/02699930341000176.
- Scotto, C. (2002). "Interacción y atribución mental: la perspectiva de la segunda persona". *Análisis Filosófico*, 22(2), pp. 135-151.
- Scherer, K.R. y Scherer, U (2011). "Assessing the Ability to Recognize Facial and Vocal Expressions of Emotion: Construction and Validation of the Emotion Recognition Index". *Journal of Nonverbal Behavior*, 35(4), pp. 305-326. DOI: 10.1007/s10919-011-0115-4.

- Smith, J. (2015). "The Phenomenology of Face-to-Face Mindreading". *Philosophy and Phenomenological Research*, 90(2), pp. 274-293. DOI: 10./phpr.12063.
- Smith, J. (2016). "Perception, emotion and value". En J. Dodd (Ed.), Art, Mind, and Narrative, Themes from the Work of Peter Goldie (pp. 121-137). Oxford: Oxford University Press.
- Spaulding, S (2015). "On Direct Social Perception". *Consciousness and Cognition*, 36, pp. 472-482. DOI:10.1016/j.concog.2015.01.003.
- Stout, R (2010). "Seeing Anger in Someone's Face". *Proceedings of the Aristotelian Society Supplementary*, 84(1), pp. 29-43. DOI: 10.1111/j.1467-8349.2010.00184.x.
- Wellman, Henry M. (1990). *The Child's Theory of Mind*. Cambridge: MIT Press.
- Wieser, M. J. y Brosch, T. (2012). "Faces in context: a review and systematization of contextual influences on affective face processing". *Frontiers in psychology*, 3(471), pp. 1-12. DOI: 10.3389/fpsyg.2012.0047.
- Woodworth, R. S. (1938). *Experimental psychology*. Nueva York: Holt.
- Zahavi, D. (2011). "Empathy and direct social perception: A phenomenological proposal". *Review of Philosophy and Psychology*, 2(3), pp. 541-558. DOI: 10.1007/s13164-011-0070-3.
- Zloteanu, M., Krumhuber, E.G.y Richardson D. C. (2018). "Detecting Genuine and Deliberate Displays of Surprise in Static and Dynamic Faces". *Frontiers in Psychology*, 9(1184), pp. 1-8. DOI: 10.3389/fpsyg.2018.01184.

Parte III. Desarrollo ontogenético y multimodalidad

Desarrollo semiótico, multimodalidad y cuerpo

FERNANDO G. RODRÍGUEZ (UAI)

Desde hace más de medio siglo existe el apotegma que defiende, para el aprendizaje de la lengua y para la adquisición del lenguaje como proceso psicológico de rango general, un curso veloz, mayormente sencillo (en proporción con las dificultades cognitivas que implica una lengua), y una esencial independencia del ambiente, relegado a hacer solo un aporte subsidiario: gatillar un mecanismo de muy refinada ingeniería y alta abstracción. La sencillez-velocidad de un lado y la complejidad de la estructura cognitiva de otro tienen como consecuencia, si se toman en conjunto, que el lenguaje se convierta en parte del acervo innato del saber humano, ya que de otra manera no sería explicable esa conciliación. A la fecha, sin embargo, la evidencia empírica a favor de este innatismo se ha mostrado más bien elusiva, y el programa de investigación científica que supo cobijar esta ambiciosa afirmación (cognitivismo fundacional, con su reconocida inspiración en la informática) ha resignado sus aspiraciones de explicar el universo psíquico, el *software* del sujeto, con omisión expresa y radical del cuerpo, el *hardware*, largamente condenado al ostracismo por cierta realizabilidad variada, disolvente de las diferencias entre sistema nervioso y sistema de cables, engramas y microchips, experiencias encarnadas y computativas. Con todo, las voces

antirreducciónistas y antifuncionalistas no siempre reivindicaron la corporeidad, sino destacaron como inadecuada la homologación entre computaciones informáticas y acción sináptica. En efecto, la desavenencia pasaba por el cerebro (Searle, 1989), no por la vitalidad del cuerpo en movimiento, capaz de ofrecer –a cuales fueran los dispositivos de procesamiento en el sujeto– una clase particular de información, clase indicada por el hecho de que el CPU en cuestión puede moverse, está dotado de dos extensiones de aptitud mayúscula “todo terreno” (manos) y una sensibilidad íntero- y propioceptiva que afecta, de un modo decisivo, la computación de representaciones *objetivas* (Gallagher, 2005).

El cognitivismo de primera generación fue mentalista a ultranza. La renovación, fruto de los cerebralistas (Churchland, 1989; Bunge, 2010), dejó nuevamente postergado al cuerpo como tal. Pero resulta manifiesto que lo que el cerebro puede elaborar y aquello que la cognición puede llegar a conocer está sesgado por las condiciones morfológicas y orgánicas del cuerpo, responsables en última instancia de lo que arriba al cerebro, el cual recibe en suministro no tan solo contenidos, sino coordenadas de intuitividad (como aquellas determinadas por la posibilidad del movimiento en el espacio). Pensamos con las categorías que el cuerpo determina: me muevo antes de hablar o despejar la incógnita de una ecuación. La marca de este condicionamiento había sido olímpicamente soslayada de las discusiones académicas de la filosofía analítica, y algunos autores de la tradición continental han sido, por esta razón, recuperados en tiempos recientes (Husserl, Merleau-Ponty). El cuerpo como formante, agente configurador del mundo por las restricciones morfomateriales que impone al trabajo de procesamiento del cerebro: tal la visión del giro corporal (Sheets-Johnstone, 2009).

En una breve primera sección, se plantea el marco de nuestro trabajo general: los puntos ciegos de la versión modular-innatista del lenguaje y su versión alternativa desde la psicología del desarrollo. En la segunda sección, se traza una cartografía esencial de la ontogénesis de la semiosis en el niño pregramatical, poniendo en primer plano que las habilidades sociocognitivas condicionan y motorizan los progresos en este terreno, particularmente las ligadas al uso del cuerpo en las interacciones con el semejante, de un modo específico las vinculadas con el gesto. En la tercera sección, se profundiza en la clase de los gestos representacionales, en cuyo seno la plasticidad semiótica del cuerpo se presta de forma iconográfica a la comunicación de contenidos de tipo referencial. El recorrido pretende mostrar la máxima importancia de reintroducir el cuerpo en los estudios sobre el desarrollo ontogenético de la semiosis *sensu lato*, y de modo específico en estudios sobre las primeras fases del lenguaje, concebido prioritariamente como *actividad*.

1. Acción semiótica y lenguaje

El modelo (psico)lingüístico hegémónico durante buena parte de la segunda mitad del siglo XX ha sido el generativista, cuya aptitud o propiedad descriptiva -en términos de cómo se procesan cognitivamente las unidades simbólicas afectadas a producir enunciados- no evaluaremos aquí, pero del cual hemos de señalar que sus apreciaciones sobre la emergencia del lenguaje parecen apenas resistir una contrastación con la evidencia empírica que, de modo creciente, alienta por cierto explicaciones de tenor evolutivo. Nada tiene que hacer en esta discusión el cuerpo, a menos que, junto con la palabra o las lenguas de señas,

se convide en ella al gesto y a la comunicación del tipo no verbal en su más rica dimensión. Naturalmente, para hacerlo hay que apartar dos postulados de la perspectiva generativista: innatismo y modularidad. Si se defiende el presupuesto de una modularidad formal, el gesto no tiene cabida en las explicaciones del lenguaje. Pero ocurre que el gesto se ha mostrado un colaborador vital en el curso normal de toda adquisición lingüística (Butcher y Goldin-Meadow, 1993; Capirci, Contaldo, Caselli y Volterra, 2005; Capirci y Volterra, 2008), y que la modularidad se rompe al asumir que las composiciones de gesto y palabra integran signos (unidades informacionales) de características muy diferentes (gestáltico el primero, articular y flexivo el segundo) en estructuras cognitivo-comunicativas solo concebibles al precio de elastizar los muros del lenguaje *puro* (McNeill, 1992, 2005; Kendon, 2004, 2007).

En relación concreta con el innatismo, la psicología ha sabido responder, de modo más amable o menos, a las presuposiciones y a las conclusiones de razonamientos con premisas asumidas demasiado pronto como verdaderas. El innatismo lingüístico tiene, por lo pronto, tres grandes objeciones que afrontar:

1. *Objeción histórica:* El viejo planteamiento dicotómico que asume tener razón por descontar que el conductismo está descaminado y afirma rotundamente que el lenguaje no puede aprenderse (no es explicable desde la experiencia) es una alternativa reducida a extremos (todo por estimulación o todo inscripto en el genoma) que no representa los estudios de psicología del desarrollo actuales. Pudo haber sido una estrategia interesante en los 60 (Tomasello, 2005), cuando precisamente el conductismo enfrentaba reparos respecto de su aptitud para explicar las funcionalidades

superiores de la vida anímica, pero aquel silogismo disyuntivo no puede aplicarse a menos que se pruebe que p o q son falsas. Hoy disponemos de un surtido espectro de variantes intermedias, toda una gama de teorías que intentan explicar el surgimiento del lenguaje, de suerte que ya no existe un diálogo polarizado (un contrapunto que había soslayado otras alternativas del medio académico europeo), sino una polifonía de perspectivas divergentes pero también convergentes.

2. *Objeción empírica:* El innatismo no termina de explicar cómo el niño pequeño, después de haber adquirido una regla determinada, comienza de pronto a cometer errores que desdicen la previa actuación correcta (conjuga “sabo”, cuando decía normalmente “sé”). Puede alegarse que quizás nunca tuvo la regla y solo repetía. Pero eso contradice la evidencia de que, mientras decía “sé”, enunciaba las formas verbales de 1.^a singular presente indicativo de verbos en “-er”, regulares e irregulares (“corro”, “como”, “puedo”), de un modo correcto, por lo que parece lo más razonable suponer que sí se hallaba en posesión de la regla del caso. Una versión mejor acomodada a este comportamiento extravagante sugiere que a una gramaticalidad de tipo defectivo sucede más tarde una segunda etapa, rígida, que proyecta las reglas más allá de sus fronteras. En este punto el niño habría entendido que la lengua es un sistema, aunque por el momento la tome por un sistema matemático. A espaldas de estos errores, la sobrerregularización supone por ello un importante avance cognitivo. La alternativa entre una estructura innata de conocimientos y un proceso dinámico nos conduce al problema siguiente.

3. *Objeción epistemológica*: Entre dos teorías que expliquen lo mismo, por principio de parsimonia de la ciencia, goza del favor natural la que comprende un menor margen de suposición (hasta que haya evidencia que se pronuncie en contrario). Sin cometer una falacia *ad ignorantiam* (pretender que la falta de pruebas oficie como elemento probatorio de la propia posición), la alternativa emergentista puede defender una mayor fidelidad a los fenómenos. Como hace tiempo ha precisado ya Manuel De Vega:

La distinción competencia-actuación es un dispositivo de inmobilización perfecto. No hay forma de falsear este tipo de teorías, cualesquiera que sean los datos obtenidos en las investigaciones. Si los sujetos obtienen pautas de rendimiento predichas, ello se lo atribuye a la competencia. Si obtienen rendimientos inesperados, se defiende la competencia a pesar de todo y se atribuyen los errores a cualquiera de los factores psíquicos que hemos mencionado [...]. ¿Cómo es posible que una competencia innata, universal y de uso general esté tan obscurecida por elementos accidentales, hasta el punto de que el sujeto actúe en contra de su propia competencia? (De Vega, 1984, pp. 464-465).

Las tres grandes objeciones mencionadas, por no abundar, nos llevan a la relativización de algunas tesis fundamentales de las aspiraciones del innatismo. ¿Por qué se ha dado como tan seguro que los niños aprenden el lenguaje rápido y sin esfuerzo? ¿Cómo se mide el esfuerzo mentado? En favor del modularismo, se ha alegado el caso de la disfasia (*SLI: Specific Language Impairment*), un déficit genético que al parecer no es, como alguna vez se creyó, tan específico de la capacidad lingüística, sino que compromete “a praxias faciales no lingüísticas, a la cognición visual y al CI global; parece ser que el procesamiento de material secuencial rápido está afectado en un niño disfásico” (López Ornat, 1999, pp. 476-477), como se aprecia

en que tenga dificultad para el procesamiento temporal auditivo de la señal de habla (para la decodificación de la secuencia de unidades de sonido con significado). De ello, sería la percepción del lenguaje lo que llevaría a una gramaticalización deficitaria, y no una gramática lesionada en su presunta biología.

Concretamente, lo específico en nuestra capacidad para el lenguaje puede ser menos una estructura o programa informático, heredado de épocas prehistóricas, que un mecanismo o predisposición (igualmente prehistórico) capaz de activarse desde competencias previas (habilidades para procesar estímulos sonoros, para compartir estados mentales o afectivos con otros, para reconocer en ellos a sujetos con intencionalidad) que deben cruzarse, fecundarse recíprocamente y posibilitar la entrada al signo y la palabra, penetrar la significación del otro y poder generarla (comprensión/ producción). Las personas afectadas con el síndrome de Asperger hablan, pero evidenciando un sentido del otro perturbado (American Psychiatric Association *et al.*, 2014). Que hablen respalda, de hecho, la idea de un lenguaje modular, salvo por la objeción de que su comunicación implica al otro de alguna manera, aun si no accede al estado mental del interlocutor. Sin cierta conexión de base con un alter ego en algún punto semejante, no habría tampoco el lazo (entiéndase: canal) por el que hacer correr los signos del lenguaje, ni motivos para hacerlo. El otro está de alguna forma siempre para posibilitar el signo y la sintaxis.

Lo que hay que discutir son pues las condiciones de la adquisición. Si el lenguaje consiste en gestionar mediante signos nuestras vidas con los otros, el manejo de rubros formales con los que se pueda elaborar complejos enunciados recursivos no es más que la meta psicológica esperable, el nivel superior de las interacciones, no el punto

de partida. Si, al contrario, cabe remitirlo a restricciones cognitivas que trabajan sinérgicamente en su emergencia y crecimiento, en el albor de la palabra no es preciso suponer una organización formal lista para operar. Ciertas claves semióticas surgidas en los intercambios cotidianos pueden explicar el paulatino devenir del acto comunicativo en habla y en habla gramatical. Toda la discusión es, pues, si el conjunto de recursos sociocognitivos del niño pequeño es accesorio o decisivo en el proceso de la adquisición lingüística.

Hay condiciones madurativas y sociales para la adquisición del lenguaje. La interpretación pragmática de las rutinas cotidianas, rutinas que empiezan a significar por suspensión de acciones, sugiriendo la totalidad de lo que ha sido interrumpido (Nelson, 1996; Rivière, 1986, 1990), es la cantera del significado. Cuando estas rutinas, que fundan sobreentendidos, habilitan la omisión de ciertos elementos cognitivos en la comunicación expresa y la actividad materna, por ejemplo, se asocia a la consecuencia, a lo contiguo o lo semejante, lo primero *significa* lo segundo. Así despunta un primer código, mezcla de lengua y gestos, de hábitos y expresiones faciales, de miradas y guiones preestablecidos, código capital, tosco, embrionario, aquel idioma primigenio, interactivo y, desde luego, intermodal, que pese a su precariedad irá alcanzando poco a poco suficiencia para hablar del mundo sin limitaciones. Se trata de lo que Bruner (1981) había señalado con su *Language Acquisition Support System*.

No hay pruebas de innatismo, sí en cambio de que un desarrollo nos conduce de no hablar a conseguir hacerlo. Y en ese desarrollo el cuerpo contribuye de maneras fundamentales (no del todo teorizadas), y no tan solo con funciones de soporte en la articulación de la palabra. Si durante mucho tiempo el cuerpo nos fue presentado como

realidad casi antitética respecto del lenguaje (este, por cierto, un miembro numerario de las cogniciones superiores, aquél, morada de pasiones y necesidades), parece no obstante poder recusarse esta separación, mínimamente en lo que toca a la *emergencia* del lenguaje. En la medida en que este vive ya en el habla pregramatical, cuyo patrón de coordenadas más elemental es una relación de cara a cara y por ende multimodal (tonos de voz, color de voz, prosodia, gesticulación, interacciones, expresividad y un largo etcétera), el cuerpo está afectado, de un modo esencial, en diversos procesos de semiosis que afluyen hacia el lenguaje y preparan su advenimiento.

Este proceso nos revela que, antes de cualquier prueba de gramaticalidad, hay en el niño un uso pertinente de palabras, solas y con apuntalamiento en otras variedades de semiosis. Este doble nivel de la palabra pregramatical y ya gramaticalizada remite a los dos niveles del lenguaje presentes en Aristóteles (*onomázein*, ‘nombrar’, distinto de *légein*, ‘decir’ ‘predicar’). Entre otras cosas, la reivindicación del rol del cuerpo en el proceso de adquisición del lenguaje debería propiciar un cuestionamiento de esta dicotomía, destacando que la predicación es ya posible para el niño en un formato originario de multimodalidad. Puede expresar, como será ilustrado en la tercera parte, protoproposiciones con la mera yuxtaposición gesto-palabra.

2. El desarrollo de los signos en la infancia temprana: algunas consideraciones

No será redundante repetir aquí ciertas definiciones para guiar en las secciones subsiguientes. Para empezar, el signo no es una entidad tan obvia como podría parecer. La ver-

sión tradicional (*aliquid pro aliquo*) es demasiado generosa y laxa; la versión de Eco (1976), según la cual es signo todo aquello con lo que puede mentirse, es al contrario demasiado estricta, deja a un costado grandes territorios a los que difícilmente podría denegárseles valor semiótico -las imágenes, la arquitectura (Nöth, 2000)-. Llamaremos "signo" a toda entidad con *vocación significante*, esto es: toda entidad cuya materia física o virtual se desdibuje como tal conforme con un ejercicio remisivo consumado más allá de sí, de una manera compartida y decodificable, por uno o más interlocutores. El desdibujamiento de la materia significante *qua* materia o realidad en y por sí, junto con la obligada remisión semiótica, no se confunden con la asociación incidental, con un acto de inteligencia, con el *insight* o la repentina captación de un lazo entre dos episodios o fenómenos, sino que implican una codificación, un compartir claves relationales con los interlocutores. El signo supone *prima facie* una articulación entre lo que se anula como cosa (para semantizarse) y aquello que de esta forma es evocado. El puño con el índice en una dirección determinada es signo, gesto de señalar, solo si el receptor u observador, ocasional o no, ve en la morfología adoptada por la mano una flecha o vector, nunca una mano sino cierta proyección imaginaria que parte del dedo y se prolonga hasta una diana intencional.

La significación surge dentro de la ontogénesis en andas de la acción y de la coseidad inerte, se emancipa de las prácticas concretas compartidas del niño y su entorno familiar y se transforma en comunicación cuando los signos son empleados con intencionalidad (Español, 2004; Rivière, 1984, 1990). Por cierto, no es preciso dedicarse a la psicología del desarrollo para comprender que el niño nace a un universo en que la significación se encuentra ya a su alrededor y plenamente activa desde antes del habla

y de la gesticulación. Pero esta apreciación capta la escena desde fuera, desde la atalaya de una tercera persona oracional y desde el consecuente modelo académico. En rigor, el ingreso del niño a los procesos de semiosis, lo mismo en el rol de receptor-intérprete que en el más exigente de emisor, debe esperar a que su cognición enlace objetos o experiencias *de manera tal que* un primer elemento auspicie a otro segundo y se anule a sí mismo fuera de su acción de remisión. La lenta marcha a este destino de sujeto capaz de semiotizar se inicia tempranamente con recursos que pueden ligarse de forma mediata con la lengua. Aunque imbricados de forma lejana, precursan las ulteriores habilidades comunicativas y, por su intermedio, el acceso a la lengua. Hacia el último trimestre de la gestación, es posible al feto discernir los sonidos vocálicos, las emisiones de varón y de mujer y la voz de la madre de modo particular (Einspieler, Prayer y Prechtl, 2012). Esta incipiente destreza auditiva permite el primer enlace con el medio extrauterino mediante el registro de las vibraciones háptico-sonoras a través de los resonadores del cuerpo materno y el sistema óseo. El feto recibe así una información que amplía su espacio ínteroceptivo y propioceptivo (Van de Carr y Lehrer, 1988).

En el nacimiento, las primeras experiencias auditivas, y las perceptivas en general, no son, como creyeran William James y Jean Piaget, un caos incoordinado y multisensorial que el niño alcanzaría a integrar de a poco. Parece probado que la percepción evoluciona desde la intersensorialidad hacia la discriminación de los canales de ingreso de información (Bremmer, Lewcowicz y Spence, 2012). La multimodalidad es la característica de la inmediata afiliación entre *inputs* que ingresan al cuerpo y al sistema cognitivo por vías diferentes. Pero no solo opera en los niveles de la sensopercepción, sino que también

participa, con un margen de responsabilidad incontestable, en el pasaje de las comunicaciones prelingüísticas a las lingüísticas, y en el de las formas lingüísticas más despojadas, la palabra aislada o pregramatical en relaciones cara a cara, hasta los cotos de modularización y autonomización de la palabra con morfosintaxis (Karmiloff-Smith, 1992). El elemento que articula esta continuidad es el cuerpo presente, su aptitud-solicitud tanto para la manifestación emocional como para alojar el mundo en sus afanes comunicativos, *a priori* y *a posteriori* de la gramaticalización.

La primera relación humana de intersubjetividad implica una opción básica de reciprocidad precomunicativa y presemiótica: una forma de entendimiento con el semejante en que este es simplemente un polo emocional con el cual compartir vivencias en una determinada sintonía afectiva. De allí que el primer vínculo realmente interactivo (dejando a un costado la respuesta neonatal de tipo imitativo, difícilmente asimilable a un involucramiento volitivo) lleve la marca estructural de un vitalismo, entendido a partir de las llamadas por Stern (2010) “formas de la vitalidad” (la experiencia integrada del movimiento, el tiempo, el espacio, la fuerza y la direccionalidad/intencionalidad -que el propio movimiento engendra-), responsables de la comprensión no conceptual de los estados de ánimo del interlocutor. Este conjunto de factores ordena, por cierto, el ida y vuelta en las interacciones bebé-adulto, y, cuando más tarde medien signos, lo harán sobre una relación originaria como añadidos que la dotarán de mayor sofisticación y nuevas posibilidades. Esa estructura primitiva es por demás multimodal, y esta multimodalidad es la matriz dentro de la que se despliega la experiencia de *ser dos uno-con-otro* en simultáneo y por el gusto que ello importa para los involucrados. Hecha de movimientos y sonidos, esta experiencia ha recibido el nombre de

“musicalidad comunicativa” (Malloch y Trevarthen, 2009), aunque siguiendo la acepción que hemos tomado para “comunicación” (la significación deliberada y para un *partenaire*), mejor sería rebautizarla “musicalidad temprana” o bien “interactiva”⁶⁷. Las variables de esta musicalidad engloban el timbre y la altura, pero no menos el pulso, el ritmo, la intensidad y la melodía en el movimiento corporal: se conjugan en ella los sonidos y su correlato kinestésico. La trama de intersubjetividad que así resulta (Español, 2008, 2014; Español, Martínez, M., Bordoni, Camarasa y Carretero, 2014; Español y Shifres, 2015; I. Martínez, 2014) modela el primer encuentro con la lengua, permitiendo al niño distinguir sonidos, seccionar la cadena indiferenciada de la verbalización adulta e ir vinculando esos segmentos con significados. Solo que el acceso a la lengua nunca es tan lineal.

¿Cómo acontece el paso entre esta reciprocidad o mutuo entendimiento y el momento en que comienza a haber interacciones significativas? El lenguaje surge de un proceso que metaboliza de forma crecientemente refinada los estímulos de entorno, que encierran en sí todo cuanto es preciso, aunque la adquisición, esto es, el aprovechamiento de los contenidos disponibles en el *input*, ocurre por grados sucesivos y en conformidad con el progreso de la cognición. La metabolización de información oral extraída a la cadena de habla del adulto se apoya en distintas competencias de tenor social. A partir de los dos meses, el bebé puede entablar con su interlocutor adulto una especie particular de relación, una experiencia lúdica de manifestación oral y movimientos, respetando turnos alternantes que han justificado el nombre de protoconversaciones

⁶⁷ Sugerencia de M. Bordoni, cuyo trabajo en el tema puede consultarse en Bordoni y Martínez, 2011; Bordoni y Español, 2018; Español, Bordoni, Martínez, M. y Carretero, 2015).

por su semejanza con la interacción verbal de los adultos. Este intercambio emocional-vital diádico, en el que el bebé se muestra participativo y reconocedor del otro como un otro al cual puede incitarse a este juego social, es considerado como la primera forma de intersubjetividad (Trevarthen, 1980). Hacia los 9 meses, la estructura de la interacción es muy profundamente conmovida al convertir su originaria condición diádica en triádica (M. Martínez, 2010). Esto acontece con base en el mecanismo de atención conjunta por el que el bebé ha desarrollado sus capacidades perceptivas respecto del *partenaire* humano, hasta captar, en el vector de la mirada del adulto, la relación de este con un objeto intencionado. El establecimiento de esta conexión permitirá al bebé empezar a coordinar su propia acción con la acción del adulto, y de esta forma entrelazar sus intenciones con las de los otros, estableciendo los patrones conductuales que van a regir definitivamente su vida social. Al detectar la intencionalidad del otro sobre los objetos, se inaugura para el niño la opción de influenciar los estados mentales de sus semejantes. Solo falta darse a entender con eficacia.

El primer signo producido por los niños parece explicarse por cierto desgajamiento a partir de la acción. La ritualización de una conducta (su desconexión de la función por la que el organismo obtiene un beneficio y su estereotipización) la convierte en señal. Como afirma Español, “un gesto se destila, entonces, de una actividad social preexistente” (2004, p. 89), desafectando la actividad de la consecución de un fin práctico.

Cuando un niño usa un gesto [...] se muestra consciente de (1) estar realizando una acción que no tiene directamente efectos en el mundo material sino en el espacio de su interpretación, es decir, en un espacio interpsicológico, y (2) de que su

gesto tendrá las consecuencias esperadas, o no, en función de la acción que esta interpretación induzca en el intérprete (Español, 2004, p. 106).

El desprendimiento de una forma y un significado a partir de una realidad monádica cerrada (cuya característica de cierre o completud precisamente se resigna al depender, en adelante, del aporte de interpretación del otro, necesario para completarse) implica como tal una metamorfosis del mero comportamiento *llano* en un comportamiento *comunicativo* (Vygotski, 1931/2000).

Si la intersubjetividad primaria está regida por *convergencia afectiva* y por la inmadurez cognitivo-corporal del bebé, la secundaria comprende no solo un conocimiento mucho más rico del entorno y del sujeto *partenaire*, sino una triangularidad dinámica donde los interlocutores se influyen *no* por la acción, *sino* por mediación semiótica. Con la atención conjunta, y sabedor de que el adulto puede interpretarlo, el bebé concreta un *pointing*, gesto de señalar, por lo común interpretado como indicación de eso que quiere, protoimperativo, o quiere compartir con otro foco de interés, protodeclarativo (Bates, Camaioni y Volterra, 1975).

¿Es verdaderamente el *pointing* la primera gesticulación deliberada? En tiempo reciente, se ha propuesto que determinados gestos, reunidos con la etiqueta de “ostensivos” (*showing*, *giving*, *placing*: gestos de mostrar, de dar, de colocar ante el otro), son las formas iniciales de la significación intencional (Clark, 2003; Rodríguez, Moreno-Núñez, Basilio y Sosa, 2015). En particular, la categoría de la ostensión parecería poder desagregarse, sin sufrir un forzamiento, en el *giving* y el *placing*, por un lado, y el *showing*, por otro, defendiendo que aquellos primeros son casos de acción semiótica y este, un signo genuino. La diferencia estriba en la noción de signo y el lugar que se conceda al

desarrollo en el proceso de la adquisición de habilidades comunicativas. Si los signos son acciones insolventes para la consecución de fines sin la mediación de otra persona, lo que distingue el empuñar una cuchara para alimentarse y en un juego de comer es por sí mismo manifiesto. Pero el discernimiento se torna complejo cuando se trata del *dar*. Es una acción, eso está claro: consiste en poner en el espacio de otro, cuando no en su mano o su regazo (si se trata del niño pequeño), un objeto determinado. No obstante, y aunque el dar tenga valor de acción instrumental, cierta actitud de parte del agente lo carga de expectativa y significación: el niño se queda mirando al otro para constatar si este se ha percatado de haber recibido algo que hacía un momento no se hallaba en su poder, y que ahora tiene por su voluntad y agencia. A la acción efectiva sucede una pausa y una expectación, compás de espera y de tensión semiótica por el completamiento de la acción, que es la interpretación del otro. Tomemos este ejemplo: me pongo en la cola para la línea de cajas del supermercado; *me ubico* a mí mismo, me sitúo físicamente donde corresponde para realizar el pago, y cuando, luego de avanzar, quede primero, detrás de ninguno, *significará*, para los otros, para mí y para el cajero, que es mi turno. Segundo ejemplo, también dentro del supermercado (tomado de Clark, 2003): pongo en el mostrador mi tarjeta de crédito o billete. En esa acción *estoy pagando*, pero al mismo tiempo ha de entenderse que aquello depositado no es solo dinero, sino *aquel dinero* con el cual quiero pagar cierta mercadería de ese local. Solo este código en común con el cajero es la señal de que puede tomar lo que he librado a esos efectos, por lo tanto, sin derecho de mi parte a reclamar que estuviera tomando algo indebido. En este punto, la acción realizada no es ya pura acción, pero no es todavía un signo acabado. Si se admite la distinción, se hace posible defender una

continuidad y una transformación que llevaría del *giving*, acción significativa, hasta el auténtico gesto de *showing* (Español, 2004), signo en sentido propio. Esta versión del desarrollo gesticulativo se remonta hasta Vygotski (1931/2000; cfr. también Bates, 1976; Werner y Kaplan, 1963), y propone que las acciones, como ya hemos discurrido, se anticipan a la significación y le prestan su forma y contenido, convertidos respectivamente en el significante y en la significación⁶⁸.

Para que haya signo es exigible, como condición “de mínima”, este desdoblamiento de significante y significación: que aquel renuncie *materialmente* a sí mismo para convertirse en un agente trascendente más allá de sí. Hay, pues, acciones, acciones semióticas (para las cuales, en paralelo con la acción instrumental, se despliega un campo de interpretación) y gestos (derivados de la suspensión de acciones prácticas o instrumentales, limitados a ser *pura significación*). Si se defiende la importancia explicativa de la observación empírica con amplitud multimodal y el abordaje desde el desarrollo ontogenético, el eslabón de las acciones decodificables parece, por cierto, la estrategia mejor avenida a describir la génesis del hecho significativo autónomo. Obviarlo es, al contrario, maquetar todo el proceso en términos de opción por sí o por no, semiosis presente/ausente, sin admitir que existan fases intermedias. Nuevamente, aquí la referencia insoslayable es Español (2004, 2007).

Como quiera que sea, el gesto deíctico es la primera expresión sígnica deliberada de tipo referencial. Por supuesto, con el gesto deíctico todavía estamos lejos del

⁶⁸ Se engloba como “significación” tanto el significado como el referente, sin entrar en meollo del tipo de si todo signo implica referencia o contenido no referencial. En uno u otro caso, al signo corresponde un correlato aludido desde el registro manifiesto.

lenguaje. Incluso en este plano, no sería posible establecer una separación entre el primate humano y los demás, dado que los simios señalan (Gómez, 2005, 2007). Lo hacen, no obstante, solo protoimperativamente, para pedir o exigir, utilizando al interlocutor como herramienta, mientras que el niño señala *también* para enseñar al otro que algo le ha llamado la atención, convidiéndole a ver un objeto que supone pueda ser de su interés, lo que implica haber entendido al otro como un semejante cognitivo-emocional capaz de interpretar su signo, pero, sobre todo, de sintetizar de una forma distinta en los actos de referencia al mundo, sin mediar interés o ventaja en el uso de signos. Lo así señalado no es entonces para el emisor, sino para los otros, y en consecuencia la atención conjunta se bifurca en dos circuitos tripartitos que deben discriminarse. De una parte, el semejante es mediación para acceder hasta el objeto (gesto protoimperativo) y se vuelve en cierta medida un elemento prescindible al alcanzar la meta; de otra parte, contra la presunción de establecer una secuencia que haría escala en el objeto para proyectarse desde allí hasta el *partenaire* como interioridad o subjetividad, sucede sin embargo que aquí no es posible, en el final, desestimar ninguno de los elementos, porque el sujeto necesita simultáneamente al objeto y al receptor para cumplir con su intención de ubicarse con otro *respecto del mundo*. En otros términos, para el gesto protodeclarativo no es posible reducir la situación ternaria a una estrategia donde sea reconocible una matriz binaria de sujeto-objeto: en ella, el tres, y, de manera peculiar, el rol del interlocutor, hace a la esencia del tipo de relación en juego.

Si en los comienzos de la intersubjetividad secundaria, definida a partir de la atención conjunta, tenemos ya una intersubjetividad de signos, hacia el final del primer año tenemos ya una *intersubjetividad semiótica multimo-*

dal, porque a los gestos iniciales se agrega el signo verbal. Desde un punto de vista no estrictamente lingüístico, el período denominado “holofrástico” debería aquí rebautizarse e integrarse al proceso que se ha iniciado ya con la gestualidad, dado que, en rigor, la aparición de la palabra no hace más que habilitar otro canal para la significación que se suma a los recursos comunicativos previos. La novedad de la palabra aislada pierde relevancia si dejamos de observarla desde el producto final, que es el lenguaje, respecto del cual oficiaría como el amanecer todavía muy lejano. La periodización *semiótica* debería en cambio destacar lo que es, sin lugar a dudas, el mayor avance cognitivo-comunicativo de esta fase, a saber, la posibilidad de componer los signos preexistentes con los nuevos, gestos con palabras. Se trata, en definitiva, de cómo ambos tipos sígnicos cooperan desentendidos de aquellas sus marcas particulares.

La importancia de las composiciones bimodales de gesto-palabra no puede minimizarse (Butcher y Goldin-Meadow, 1993; Capirci *et al.*, 2005; Capirci, Iverson, Pizzuto y Volterra, 1996; Capirci y Volterra, 2008; Goldin-Meadow y Brentari, 2017; Volterra, Caselli, Capirci y Pizzuto, 2005). Allí, el gesto ocupa el sitio que, más tarde, en el período de dos palabras (siempre según periodización lingüística), corresponderá a un segundo signo de tipo verbal, como si aquel asumiera la función de preparar el territorio para el desembarco posterior de otro elemento léxico, dando comienzo a un encadenamiento que iría perfilando, por creciente acumulación, el arribo de reglas de composición morfosintáctica, la ordenación serial de las palabras sobre el eje natural -lineal- de la emisión, y que emanciparía a la lengua de todo apuntalamiento en signos de otra laya. Las composiciones bimodales recorren estas tres clases presumiblemente sucesivas (Volterra, Caselli, Capirci y Pizzuto,

2005): originariamente, ambos tipos de signo se encuentran consustanciados como portadores del mismo significado (composición equivalente: gesto y palabras dicen exactamente lo mismo, como en “esto” + *pointing* correspondiente); luego aparecen formas mixtas donde, en pos del mismo referente, los dos signos aportan información diversa (composición complementaria: “agua” + *pointing*, donde está claro que la información verbal, categorial, no coincide con la del gesto, de tenor indexical); por último, los signos se distribuyen entre dos contenidos semánticos plenamente diferenciales, y se generan las primeras construcciones de carácter (proto)proposicional o de predicción reconocible (composición suplementaria, en la que el niño verbaliza “yo” y señala hacia el objeto que desea, o se lleva la mano al pecho en actitud deíctica, indicando *quién*, y verbaliza el nombre del objeto o la acción que realiza).

El interés de estas composiciones consiste en mostrar cómo ya ciertas estructuras expresivo-comunicativas se organizan, empleando dos signos de extracción diversa, para transmitir los contenidos de dos representaciones mentales distintas (relación agente-objeto, agente-acción, etc.), antes de que al sistema cognitivo del sujeto sean posibles configuraciones de única modalidad. La solidaridad expresiva entre la modalidad gestual y la verbal hay que entenderla no tan solo como un paso previo en la constitución del lenguaje en el niño, lo que conlleva imaginarla condenada a desaparecer luego de esta contribución, sino como el estándar de la comunicación humana en condiciones naturales (esto significa: socioculturales) a lo largo de la vida inevitablemente compartida del sujeto, todavía y siempre ordenada, incluso en tiempos de redes sociales y virtualidad, por relaciones cara a cara (cuerpo a cuerpo) en las que prevalece el contacto multimodal. La integración gesto-palabra, explorada por Kendon y McNeill de una

forma pionera, conforma la regla de los intercambios sígnicos en la comunidad humana. El abordaje bi/multimodal ofrece un mirador distinto para investigar el desarrollo ontogenético del lenguaje, induce a entenderlo como menos modular, menos veloz de lo que pudo parecer, y lo piensa como emergente de una praxis comunicativa en la que el niño se vale de todos sus recursos para involucrarse con el semejante (lo que comprende la utilización del cuerpo y la capacidad de *leer* en otros cuerpos). La significación, así asumida, está ligada a un *quantum semiótico* vehiculado por muy distintos andariveles.

3. El mundo a través del cuerpo: los gestos representacionales

Entre sus muchas funcionalidades (expresiva, conativa, fática y demás), la palabra habla del mundo, tanto como primeramente el gesto, y acaso mejor que el gesto. Entre el gesto deíctico en particular (categoría que engloba al *pointing* y a todas las variedades antes mencionadas), cuya función de mínima es indicar dónde, y la palabra, que en su acción nominativa nos descubre el *qué*, existe evidente superioridad informativa de parte de la segunda. Todo ente empírico tiene un lugar, pero todos los entes son de especies y géneros diferentes. Asimismo, la facultad de nombrar de la palabra se agencia a través del nombre como tal, por regla arbitrario, y a través de la onomatopeya, que disuelve esta regla en favor de un signo imitativo, motivado por los rasgos de su referente. “Pato” y “cuac-cuac” nombran lo mismo, y el niño pequeño suele usar sonidos isomórficos para hacerse entender, tanto como el adulto para ser más transparente al niño (este es un aspecto del llamado “maternés” que no suele tomarse en consideración). La alternativa para referir a un objeto del mundo por

vía de la mimesis también se encuentra en la modalidad gestual, tanto en el niño pregramatical como en hablantes competentes. Por razones obvias, el niño de escaso léxico se ve forzado a optar por la variante evocativa corporal.

En esta sección son presentados episodios de gestos representacionales obtenidos de un *corpus* de observaciones efectuadas para la investigación de habilidades comunicativas durante el período de dos palabras. Se trata de casos que ilustran las ricas posibilidades comunicativas que parten de la inventiva con que el niño puede, a falta de palabras, lograr transmitir sus intenciones expresivas. A los fines de este texto, los detalles de tipo metodológico son accesorios y pueden consultarse en otro sitio (Rodríguez y Español, 2016). La plasticidad del cuerpo, tan solícito para salvar dificultades de limitación en el vocabulario, revela altos márgenes de potencial creativo –una creatividad que se define por un ejercicio que coordina una elección y una reproducción, y un compromiso cognitivo acaso superior al que se encuentra en la onomatopeya, que por su parte no puede escapar a las imitaciones de sonido–. En efecto, la reproducción mimética implica un primer momento de recorte selectivo entre los atributos de aquello a significar y, luego, la actividad-signo a realizarse, pero que ha de tener presencia, como factor limitante, en la elección del atributo a ser representado, so riesgo patente de que la intención semiótica no pase del conato. Muy en concreto, se puede palpar en la elección del rasgo por representar la condición parcial *sive* aspectual del signo en la versión peirceana⁶⁹ (algo que no es, por cierto, igual de transparente en el signo lingüístico).

⁶⁹ “Algo que está para alguien en lugar de otra cosa en un cierto aspecto o aptitud [capacity]” (Peirce, C.P. 2.228, (1931-1935; 1958).

Por oposición al deíctico, cuyo destino es reportar un único rasgo invariable del objeto, la presencia del objeto o el objeto mismo con su localización, los gestos representacionales funcionan de modo selectivo; al mismo tiempo, esta característica es diferencial de la elección que realiza el hablante, siguiendo a Saussure, en el eje paradigmático de la emisión. Mientras que aquí el sujeto escoge entre las convenciones previas de la lengua, en la corporeidad mimética del individuo sucede que la amplitud o libertad para escoger del mundo *qué* y *cómo* representar solo está limitada por su propia habilidad de atención perceptiva y extracción, luego para adaptar plásticamente el cuerpo a aquél rasgo seleccionado, destreza bilateralmente imbricada con las decisiones de recorte.

No corresponde a los gestos representacionales la exclusividad de poder indicar la ausencia de un objeto. También los deícticos pueden llevar a cabo esta función (Liszkowski, Carpenter y Tomasello, 2007; F. G. Rodríguez, 2017). Pero, por una parte, ello se cumple solo en modo ocasional, y derivado del empleo con objeto presente, mientras que la gestualidad representacional, por el contrario, se efectúa muy rara vez ante el objeto y más bien acontece para conjurar lo ausente. Como ha sido puesto de relieve acerca de la asociación entre humo y fuego, aquel es signo de este *solo si* este cae por fuera de mi percepción (Eco, 1973, 1976). Algunos signos necesitan que el sitio del referente esté vacío, y por norma el gesto representacional permite al sujeto poner ante el receptor los más diversos referentes inactuales a través de sus características de forma y de función.

A la edad de 1;7 (13)⁷⁰, B se pone en cuatro patas acercándose a unos dinosaurios de juguete, a los que ladra con vehemencia (gesto representacional + onomatopeya). Su “guau” es juego compartido: B se expone al otro y hace un papel dentro de su propia ficción ante/para el adulto. Todo su cuerpo está lúdicamente acomodado a figurar la silueta y el comportamiento del referente animal. En la misma sesión, M se golpea contra el piso. B lanza en represalia un golpe al piso y verbaliza algo impreciso, a mitad de camino de “mano” y de “malo” (o bien nombra el agente con el que efectúa la represalia, o bien, puesto a juzgar, está calificando la moral del piso, hábito ante cada tropiezo de B con los muebles, sancionando, por ejemplo, “silla mala”). Cualquiera fuere la palabra, “mano” o “malo”, se acopla con la gestualidad que reproduce isomórficamente la acción de pegar.

En 2;4 (10), jugando con R a una pelea de Bakuganes (bolitas magnéticas, de inspiración televisiva, que dejan aparecer un monstruo diminuto), exclama “¡Ya!”. Luego de hacer un gesto de victoria elevando los brazos, mueve un puño cerrado en dirección a R. La actitud amenazante, que concuerda con la tónica del juego, sumada a la mueca hostil con el ceño fruncido, reproduce la actitud beligerante que los personajes de ficción del caso suelen practicar.

En la primera viñeta, la mimética del gesto ha recogido la forma del golpe sin llevar a cabo la acción efectiva: el manotón se da en el aire, mirando a la madre, y como si se controlara que la acción retaliativa corresponde a lo que sucedió. En la segunda, el gesto abarca un movimiento de las manos y, a la vez, una expresión facial-gestual adoptada *ex profeso*. Ambos ejemplos, al comienzo y al final de la etapa estudiada, ofrecen un significado iconográfico de sesgo literal. Con el correr de las observaciones, sin embargo, los gestos representacionales se van despegando de su

⁷⁰ Respetamos la convención piagetiana [año; mes (día)] y la siguiente notación: B/niño, M/madre, P/padre, R/niñera.

referente, aludiéndolo de modo tangencial y precisándolo a través de asociaciones por contigüidad (pauta espacial), por sucesión (temporal) o por causalidad.

En 2;0 (14), B está bebiendo. Cuando acaba, extiende el vaso a P, quien lo rechaza. Para que comprenda su acción, eleva el vaso nuevamente a la boca y bebe en falso, dejando entender que ya no queda contenido. P le pregunta: “¿Qué?”, y B coloca el vaso hacia abajo y lo sacude, señalando, con un tono de interrogación: “¿No má_?” /¿no hay más?/. En el beber sin contenido se puede apreciar una gestualidad representacional en que la acción remite a eso faltante.

En 2;2 (10), P y B se encuentran boca arriba echados en el suelo. P inquieta, viendo hacia el plafón: “¿Esa es la luz?”. B: “Sí” (con un gesto de *pointing*). P: “¿Cómo se apaga?”. B dice: “Así”, y comprime los párpados para mostrar lo que es la oscuridad, para ilustrar en qué consiste la luz apagada. Es un ejemplo sorprendente de gesto representacional para indicar la consecuencia de apagar la luz. B no parece comprender del todo el sentido de la pregunta, a la que habría debido responder mostrando qué procedimiento se realiza para oscurecer la habitación (podría haber indicado con un *pointing* el interruptor), pero ha captado que se trata de *luz apagada* y responde exhibiendo en qué consiste la ausencia de luz (figura metonímica en modalidad gestual). Es de hecho una composición suplementaria, porque el gesto y la palabra se componen de una forma en la que cada cual aporta de su cuño un contenido de significado diferente: la palabra como un determinativo que señala un *cómo* (“Así”), el gesto aportando aquel valor preciso, especificativo, del “así”. Las dos partes de la expresión se solicitan para otorgar coherencia a la respuesta.

Estos ejemplos despegan la representación del duplicado literal. Lo figurado aporta un signo que el otro debe entender como un pretexto para suscitar en él una (distinta) representación mental relacionada. La forma del referente se halla desplazada. Cerrar los ojos no señala aquí *dormir* o *sueño* (que implican ojos cerrados necesariamente), sino algo que por sí mismo no se excluye con la

condición de ojos abiertos. B ha utilizado ese significante para sugerir otro significado excéntrico. Esto revela hasta qué punto es el manejo de las coordenadas prácticas lo que faculta al agente emisor saber qué signos puede emplear, fuera de las rutinas donde son comunes, en determinadas nuevas situaciones. Y es el contexto lo que pone al receptor en situación de poder entender lo formulado, dar con las claves para este “significado ocasional”. De una manera semejante, el gesto de beber solo podía apuntar, en su contexto, a eso faltante (líquido), a comportarse informativamente (‘Ya no queda’) o conativamente (“Serví más, quiero beber”). La situación aporta el marco del que el niño extrae no solo las primeras significaciones estables, sino también la oportunidad de flexibilizar sus usos.

Una de tantas peculiaridades de este gesto es que conviven en su seno diferentes posibilidades descriptivas, tantas cuantas tolere el límite de la creación. Cada característica seleccionable del modelo ha de multiplicarse por el número de posibilidades sínrgicas con las que pueda ser representado (la selección de un rasgo implica soslayar otros rasgos posibles igualmente válidos), pero existe también la multiplicidad mimética, solo acotada por la habilidad del cuerpo para la metamorfosis. Ciertas características pueden autoimponerse por saliencia perceptiva (Kendon, 2004), pero en algunos casos lo determinante es la función (Nelson, 1996). Función a veces natural de los objetos, como cuando en 1;8 (30) B usó las dos manos, horizontalmente colocadas, para sugerir la noción de tambor, una en el rol de parche y la otra palmoteando encima; función otras veces derivada de rutinas familiares específicas, como cuando en 1;10 (2) B colocó en su sien una pistola de juguete (máximo estupor de los presentes), porque había captado el parecido con el secador de pelo. En cualquier caso, función o saliencia perceptiva, la iconografía

gestual supone las tareas de recortar y de emular *activamente* a partir del motor de una imaginación analogante. Si la palabra es arbitraria, y en esta arbitrariedad reside la coerción de utilizar significantes y significados conforme con prescripciones culturales, el recurso a la figuración corpórea ofrece al niño una vasta avenida de creatividad sin correlato en esta etapa de limitación verbal. Al producir significantes motivados, “levantados” de las cosas mismas, el niño transpone la frontera de su mucho o poco léxico y puede significar, dentro del horizonte de su propia destreza expresiva, sus talentos como imitador o su elo-
cuencia corporal.

Así, tenemos que, junto con la palabra, este tipo de gesto logra separar la comunicación del anclaje de tiempo-espacio, pero, a diferencia de ella, no se apoya en un nexo arbitrario consabido, sino en cierta habilidad especular para reproducir, reflejo del objeto empírico o abstracto que es, no obstante, *recreación* del mismo, una asimilación de lo exterior al potencial semiótico del niño. Este pone de sí algo ausente en el objeto, el cuerpo mismo como agente-esquematizador del contenido por representar. En la literatura filosófica no es inusual hallar que el simbolismo humano escapa del contexto témporo-espacial más inmediato por efecto del signo lingüístico. La gesticulación representacional impugna esta exclusividad: puede haber libertad de significación sin arbitrariedad y desde la motivación que relaciona el referente y el significante. Ocurre entonces que esta opción fuerza al sujeto a asumir un mayor protagonismo. Hay que inventar, porque la copia incluye la elección de qué copiar, qué aspecto del objeto destacar para hacerlo reconocible a los destinatarios del mensaje y, finalmente, no por eso menos importante, qué partes del cuerpo hacer intervenir. En lo verbal, hasta no hallarse el niño en posición de componer palabras,

la creatividad es una posibilidad exigua; en la modalidad gestual ella tiene cabida incluso empleando un solo signo (cerrando los ojos para sugerir oscuridad). La significación creativa surge con el gesto antes de que el lenguaje la eleve a niveles superiores, ciertamente inalcanzables para la expresividad del cuerpo (las teorías científicas no parecen posibles sino bajo vigilancia estricta sobre el grado de rigor lógico-descriptivo de sus contenidos). La palabra preexistente al sujeto, es de los otros, circula por la cultura, pero este tipo de gesto iconográfico puede acuñar una forma *sui generis* de referir. En cierto modo es posible afirmar que la palabra se aprende pasivamente (“en cierto modo” sirve para protegernos contra la objeción de que no existe aprendizaje sin un compromiso activo de parte de los sujetos), mientras que en la gesticulación figurativa hay que *ser como* o *hacer como* los objetos, y en ello el sujeto está forzado a un involucramiento que exige cierta inventiva. Los gestos representacionales son, con la palabra, un signo destinado a la emancipación del entorno inmediato, pero un signo que -en contraste con las unidades fonémáticas y léxicas, con las reglas gramaticales que los niños deben, trabajosamente, ir adquiriendo en un marco de exposición, repetición y prueba- puede significar sin otras restricciones que las del sujeto. El gesto representacional no nombra, si por nombre ha de entenderse a la acción de mentar al otro sin ninguna pista imaginaria. Más que eso, incluye al mundo entre los signos, *lo dice* en su propio idioma, idioma-mundo en el que los objetos no son bautizados caprichosamente, mundo de *Gestalten* y de propiedades que se contagian al cuerpo y constituyen un mensaje.

No se plantea aquí un antagonismo entre palabra y gesto, oralidad abstracta y fisicalidad concreta. Más bien, se trata de que ciertos signos capturan el mundo en códigos preestablecidos y otros se apoyan directamente en él.

Las onomatopeyas, antes aludidas, absorben del mundo los sonidos que objetivamente se encuentran en él, de suerte que las distintas lenguas emplean su servicio para referir *volando al ras* de los hechos concretos. En ambos casos, la onomatopeya y el gesto representacional, la forma es *también* el contenido: ella contiene al referente en la acción misma de significarlo.

La significación creativa recuerda eminentemente el mecanismo suspensivo postulado por Rivière (Rivière y Sotillo, 1999; cfr. también Español, 2003), solo que en su propuesta se trata de suspender acciones y usar los fragmentos para sugerir la idea del todo o de la porción omitida (el niño sopla el pabilo de la vela para pedir que la enciendan y poder, acto seguido, jugar a apagarla). La gesticulación representacional incumbe, sin embargo, a signos que no son extractos de rutinas o episodios singulares, sino, en todo caso, suspensiones quietas, extracciones realizadas a partir de los objetos *como tales* que, al ser significadas, obran en sus veces.

Con todo, ¿no ha suspendido el perro, por ventura, el cuadro general dentro del que tiene lugar su vuelta a la manzana cuando, la correa colgando de los dientes, va en búsqueda del amo o se sienta a esperar junto a la puerta?; ¿no ha conseguido así él también una segmentación de la rutina de paseo, no ha consagrado como signo la acción recortada? Si es una acción, y es un mensaje, la significación creativa en el sentido aislado por Rivière no es patrimonio del sujeto humano. Por ende, la creatividad para escoger un rasgo del objeto o desprender una fracción de la secuencia comportamental implica menos diferencias interespecíficas que el hecho de que el niño tope, en este punto, con la posibilidad de figurar el mundo *con su cuerpo*. Dado que el perro puede segmentar y hacer valer con calidad de signo una conducta, no es la segmentación,

sino el descubrimiento de las *posibilidades del representar* lo que implica un progreso cognitivo en el signo representacional y nos deja arriesgar una frontera entre el hombre y el resto de las especies. Luego del cuerpo como anclaje del sí mismo y de la exploración del movimiento y del espacio circundante por su medio, el niño le ha encontrado una veta desconocida.

Antes el cuerpo era solo herramienta, el medio de la acción y el movimiento, el asiento de sensaciones y experiencias, el sensor del límite entre lo exterior y lo interior. El cuerpo era el sí mismo, las formas rudimentarias de vinculación con otros (las formas de la vitalidad) más un elenco de capacidades prácticas y comunicativas de diversa clase. Ahora se ofrece al niño como todo un nuevo capital. Usar el cuerpo para la reproducción/recreación del referente es un hallazgo que libera posibilidades comunicativas por encima de las manifestaciones expresivas espontáneas y de la contigüidad de la deixis. Al asumir esta variante síncofigurativa el cuerpo se revela mucho más versátil semióticamente que lo que había demostrado con la sola indicación del referente. Al reflejar la cosa (*mimesis*), el cuerpo advierte su capacidad de *ser también* el mundo, de cobrar sus formas y de introyectarlo, de llegar también a proyectarlo nuevamente al exterior por vía de imaginárselo en plasmaciones físicas concretas. Pero para que el cuerpo aloje al referente hay que poner a este en *idioma corporal*, llevar el mundo al cuerpo. Mundo corporeizado, entonces, por virtud de la ductilidad para el remedio físico de formas y de movimientos de la realidad.

Esta imaginación del mundo al recurrir a la corporeidad del gesto despliega ante el niño un abanico de escenarios y de habilidades con las que expresarse de manera inédita. La representación mediante el cuerpo permite entrever un lado oculto de las competencias físicas y

comunicativas que, al cruzarse con las manifestaciones de empatía y de musicalidad temprana, alumbrará un espacio para que se desarrolle, sobre la base del ritmo compartido y de la exteriorización de estados afectivos, una elo-cuencia diferente, sin palabras, pero igualmente solvente para *hablar* del mundo, para modelarlo desde la blandura artística de la corporeidad. De esta manera, el ejercicio rítmico infantil originario recibe de la variante del gesto representacional una promesa de llegar plásticamente hasta las cimas de la imitación y la emoción conjuntas, hasta los pasos de ballet mediante los que el cuerpo es la arclilla maleable capaz de volverse un cisne, un cascanueces, viento, Don Quijote. En síntesis, la representación icónica a través del gesto transfigura aquellas condiciones fenomenológicas fundamentales del ser-en-el-mundo y ser-con-otros (al decir de Heidegger) de una manera camaleónica para ser como-el-mundo y a la vez como-los-otros.

4. Conclusiones generales

En nuestras dos primeras secciones, se quiso esbozar de qué manera la perspectiva cognitivista en general, y el lenguaje de modo particular, condicionado a ser un sistema de símbolos abstractos, tenían dificultades que, con la inclusión del cuerpo como una variable necesaria, se ajustan mejor a los progresos que el niño lleva *efectivamente* a cabo en su camino a la gramática. Una brevíssima recapitulación de los problemas del enfoque modular-innatista continuó en la exposición de los hechos establecidos con mayoritaria aceptación por la comunidad de los psicólogos del desarrollo, en específico de la infancia temprana, mostrando como alternativa una versión integrativa de capacidades previas del sujeto que conducirían, no sin

esfuerzo, hasta el lenguaje. La psicología de este período incluye enfoques diversos que, sin poder envolverse bajo un rótulo unitario, comparten ideas de dinamismo y de conflicto cognitivo, de equilibración y avances no lineales, de retroalimentación entre semiosis y otras competencias, con afán de recobrar, para la discusión sobre el lenguaje, otros tipos de signos, la semántica como factor central, signos lingüísticos aislados o ligados por acoplamiento (por ende, pregramaticales), y aquellos niveles sociocognitivos que se habían tenido, en algún tiempo, por contribuyentes periféricos. Las voces contrarias hablaban desde los márgenes del caudaloso delta cavado por el cognitivismo, fiel al modelo computacional-funcionalista. Desde este otro lugar, renovador (sin ser exactamente nuevo), que cabe a cierta psicología, el lenguaje puede contemplarse como un resultado mucho más que como un punto de partida, como un efecto obtenido a partir de la commixtión de habilidades previas recíprocamente actuantes y fecundas (tal la noción de emergentismo, al menos en psicología).

El despliegue apretado de los hitos principales en el ingreso del niño a la semiosis *comunicativa* (recordemos: significación intencionada para un otro concebido como válido interlocutor) recaló con un mayor detalle en la valoración del gesto, en la importancia medular que se le ha concedido en los últimos tiempos. Hemos considerado aquellas condiciones que lo hacen posible, particularmente la atención conjunta como un punto inflexional para entender cómo los niños *dirigen* sus signos a otros que han reconocido como agentes de experiencias marcadas por la intención, entes *como ellos* proyectados intelectualmente y volitivamente sobre los objetos del entorno.

Por la vía del gesto y con la aspiración de poner de relieve el gran protagonismo que el cuerpo ha adquirido en los estudios filo-psicológicos sobre la mente, en la última

sección se ahondó en el gesto icónico o representacional, normalmente presente en la investigación empírica, clasificado y contabilizado en los análisis cuantitativos, pero en igual medida postergado en lo tocante a evaluaciones integrales, teóricas, de lo que implica en el mapa de los progresos semióticos del niño (como si su naturaleza revisiera poca gravedad en la ontogénesis y el crecimiento de la comunicación humana). Por el contrario, hemos querido defender su valor específico de gesto con características particulares. Puesto al nivel de la palabra y de los demás gestos, se revela como aquel espacio, dentro del universo de la significación, en el que el cuerpo adquiere máxima presencia. En concreto, la significación corporeizada en el gesto representacional plantea una relación entre el mundo y el yo que no es paragonable con ninguna otra variante (con la salvedad hecha de la onomatopeya, a la que -hemos argumentado- supera por grados de libertad). La combinación de selección-reproducción iconográfica y corporeidad lo vuelven el único signo de esta condición, situado entre la palabra –con idéntica aptitud referencial– y la expresividad no comunicativa, inclusive asemántica, del cuerpo afectado al *movimiento artístico* –en alusión a la capacidad de transmisión/ suscitación de estados afectivos que tiene la danza, por ejemplo, para la que no cabe, quizás, pensar *stricto sensu* en términos de significación (F. G. Rodríguez, 2017)–.

El recorrido ha pretendido, a partir de consideraciones teóricas más generales y luego pasando por la discusión de observaciones de detalle, llevar interés a la inclusión del cuerpo en la ontogénesis de la semiosis. Ha sido el cometido mostrar cómo el cuerpo no puede excusarse de los estudios del desarrollo comunicativo, y, por lo tanto, de las investigaciones cognitivas en su más variado espectro, y asimismo destacar que por su medio entran al escenario

la multimodalidad, la dimensión semántica en sentido lato y la comprensión del otro en calidad de estratos decisivos para el planteamiento del lenguaje concebido, humboldteanamente, como actividad (*enérgeia*)⁷¹.

Referencias bibliográficas

- American Psychiatric Association., Kupfer, D. J., Regier, D. A., Arango López, C., Ayuso-Mateos, J. L., Vieta Pasqual, E., y Bagney Lifante, A. (2014). *DSM-V: Manual diagnóstico y estadístico de los trastornos mentales* (5.a ed.). Madrid: Editorial Médica Panamericana.
- Bates, E. (1976). *Language and Context: The Acquisition of Pragmatics*. Nueva York: Academic Press.
- Bates, E., Camaioni, L., y Volterra, V. (1975). "The acquisition of performatives prior to speech". *Merrill-Palmer Quarterly of Behavior and Development*, 21(3), pp. 205-226.
- Bordoni, M., y Español, S. (2018). "Antecedentes del juego con las formas de la vitalidad al inicio del segundo año de vida". En N. Alessandroni y M. I. Burcet (Eds.), *La experiencia musical. Investigación, interpretación y prácticas educativas. Actas del 13.º Encuentro de Ciencias Cognitivas de la Música* (pp. 76-89). Buenos Aires: SACCOM.
- Bordoni, M. y Martínez, I. (2011). Imitación mutua y juego musical. *Psicología del desarrollo*, 1(2), pp. 69-78.
- Bruner, J. (1981). "The social context of language acquisition". *Language and Communication*, 1, pp. 155-178.

⁷¹ El autor agradece la lectura y las sugerencias de la Dra. Mariana Bordoni y la prof. Bárbara Mouzo. De esta, muy especialmente, su buena disposición para polemizar.

- Bunge, M. (2010). *Matter and Mind: A Philosophical Inquiry*. Dordrecht: Springer.
- Butcher, C., y Goldin-Meadow, S. (1993). "From one spoken word to two: Exploring the changing nature of gesture". Paper presented at the Biennial Meeting of the Society for Research in Child Development (60th, New Orleans, LA, March 25-28, 1993). Recuperado de <https://goo.gl/RLGNfs>.
- Capirci, O., Contaldo, A., Caselli, M. y Volterra, V. (2005). "From actions to language through gesture: A longitudinal perspective". *Gesture*, 5, pp. 155-177.
- Capirci, O., Iverson, J., Pizzuto, E. y Volterra, V. (1996). "Gestures and words during the transitions to two-word speech". *Journal of Child Language*, 23, pp. 645-673.
- Capirci, O. y Volterra, V. (2008). "Gesture and speech. The emergence and development of a strong changing partnership". *Gesture*, 8(1), pp. 22-44.
- Churchland, P.M. (1984). *Matter and Consciousness. A Contemporary Introduction to the Philosophy of Mind*. Cambridge: MIT Press/ Bradford Books.
- Clark, H. (2003). "Pointing and placing". En S. Kita (Ed.), *Pointing. Where Language, Culture, and Cognition Meet* (pp. 243-268). Hillsdale: Erlbaum.
- De Vega, M. (1984). *Introducción a la psicología cognitiva*. Madrid: Alianza Editorial.
- Eco, U. (1973). *Signo*: Barcelona: Labor.
- Eco, U. (1976). *Tratado de semiótica general*. Barcelona: Lumen (2000).
- Einspieler, C., Prayer, D. y Prechtl, H. R. (2012). *Fetal Behaviour: A Neurodevelopmental Approach*. Londres: MacKeith Press.
- Español, S. (2003). "La suspensión como mecanismo de creación semiótica". *Estudios de Psicología*, 24(3), pp. 261-275.

- Español, S. (2004). *Cómo hacer cosas sin palabras*. Madrid: Machado.
- Español, S. (2007). "Time and movement in symbol formation". En Valsiner y A. Rosa (Eds.). *The Cambridge Handbook of Sociocultural Psychology*, (pp. 238-255). Nueva York: Cambridge University Press.
- Español, S. (2010). "Performances en la infancia: cuando el habla parece música, danza y poesía". *Epistemus. Revista de la Sociedad Argentina para las Ciencias Cognitivas de la Música*, 1, pp. 57-95.
- Español, S., Bordoni, M., Martínez, M y Carretero, S. (2015). "Forms of vitality play and symbolic play during the third year of life". *Infant Behavior and Development*, 40, pp. 242-252.
- Español, S., Martínez, M., Bordoni, M., Camarasa, R., y Carretero, S. (2014). "Forms of vitality play in infancy". *Integrative Psychological and Behavioral Science*, 48(4), pp. 479-502.
- Español, S. y Shifres, F. (2015). "The artistic infant directed performance: A mycroanalysis of the adult movements and sounds". *Integrative Psychological and Behavioral Science*, 49(3), pp. 371-397.
- Gallagher, S. (2005). *How the Body Shapes the Mind*. Nueva York: Oxford University Press.
- Goldin-Meadow, S. y Brentari, D. (2017). "Gesture, sign, and language: The coming of age of sign language and gesture studies". *Behavioral and Brain Sciences*, 40. DOI: 10.1017/S0140525X15001247.
- Gómez, J. C. (2005). "Joint attention and the notion of subject: Insights from apes, normal children and children with autism". En N. Elian, C. Hoerl, T. McCormack y J. Roessler (Eds.), *Joint Attention: Communication and Other Minds. Issues in the Philosophy and Psychology*. Oxford: Oxford University Press.

- Gómez, J. C. (2007). *El desarrollo de la mente en los simios, los monos y los niños*. Madrid: Ediciones Morata.
- Karmiloff-Smith. A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge Mass.: MIT Press/ Bradford Books.
- Kendon, A. (2004). *Gesture. Visual Action as Utterance*. Cambridge: Cambridge University Press.
- Kendon, A. (2007). "An agenda for gesture studies". *Semiotic Review of Books*, 7 (3), pp. 8-12.
- Liszkowski, U., Carpenter, M., y Tomasello, M. (2007). "Pointing out new news, old news, and absent referents at 12 months of age". *Developmental Science*, 10 (2), F1-F7.
- López Ornat, S. (1999). "La adquisición del lenguaje. Nuevas perspectivas". En Cuetos, F. y M. De Vega, M., *Psicolinguística del español* (pp. 469-533). Madrid: Trotta.
- Malloch, S. y Trevarthen, C. (2009). *Communicative Musicality: Exploring the Basis of Human Companionship*. Oxford: Oxford UniversitiyPress.
- Martínez, I. (2014). "La base corporeizada del significado musical". En S. Español, *Psicología de la música y del desarrollo. Una exploración interdisciplinaria sobre la musicalidad humana* (pp. 71-110). Buenos Aires: Paidós.
- Martínez, M. (2010). *De la intersubjetividad primaria a la secundaria: ¿qué percibe el bebé cuando mira al adulto?* Tesis de Maestría (inédita). Facultad Latinoamericana de Ciencias Sociales. Universidad Autónoma de Madrid.
- McNeill, D. (1992). *Hand and Mind. What Gestures Reveal about Thought*. Chicago y Londres: University of Chicago Press.
- McNeill, D. (2005). *Gesture and Thought*. Chicago y Londres: University of Chicago Press.

- Nelson, K. (1985). *Making Sense: The Acquisition of Shared Meaning*. Nueva York: Academic.
- Nöth, W. (2000). "Umberto Eco's semiotic threshold". *Sign System Studies*, 28, pp. 49-61.
- Peirce, Ch. S. (1931-1935; 1958). *Collected Papers of Charles Sanders Peirce*. Vols. 1-6. ed. por C. Hartshorne y P. Weiss (1931-1935) y Vols. 7-8 ed. por A. W. Burks (1958). Cambridge (MA): Belknap Press.
- Rivière, Á. (1984). "Acción e interacción en el origen del símbolo". En M. Belinchón, A. Rosa, M. Sotillo y I. Marichalar (Eds.), *Ángel Rivière. Obras escogidas*. Vol. II (pp. 77-108). Madrid: Panamericana, 2003.
- Rivière, Á. (1986). "Interacción precoz. Una perspectiva vygotskiana a partir de los esquemas de Piaget". En M. Belinchón, A. Rosa, M. Sotillo y I. Marichalar (Eds.), *Ángel Rivière. Obras Escogidas*. Vol. II (pp. 109-142). Madrid: Panamericana, 2003.
- Rivière, Á. (1990). "Origen y desarrollo de la función simbólica en el niño". En J. Palacios, A. Marchesi y C. Coll (Eds.), *Desarrollo psicológico y educación* (pp. 113-130). Madrid: Alianza.
- Rivière, Á. y Sotillo, M. (1999). "Comunicazione, sospensione e semiosi umana: le origini della pratica e della comprensione interpersonali". *Metis. Ricerche di Sociología e Psicologia della Comunicazione*, 1, pp. 45-76. Versión castellana en M. Belinchón, A. Rosa, M. Sotillo y I. Marichalar (eds.). *Ángel Rivière. Obras escogidas*. Vol III (pp. 181-201). Madrid: Panamericana, 2003.
- Rodríguez, C., Moreno-Núñez, A., Basilio, M. y Sosa, N. (2015). "Ostensive gestures come first: their role in the beginning of shared reference". *Cognitive Development*, 36, pp. 142-149.

- Rodríguez, F. G. (2017). *La comunicación humana antes de la gramática* (tesis doctoral inédita). Universidad de Buenos Aires, Buenos Aires.
- Rodríguez, F. G. y Español, S. (2016). "Bimodal compositions of gesture and vocalization at the beginning of verbal communication". *Infancia y Aprendizaje*, 39(4), pp. 661-693.
- Searle, J. (1989). "Mentes y cerebros sin programas". En E. Rabossi (Ed.) (1995). *Filosofía de la mente y ciencia cognitiva* (pp. 413-443). Buenos Aires: Paidós.
- Sheets-Johnstone, M. (2009). *The Corporeal Turn. An Interdisciplinary Reader*. Exeter: Imprint Academic.
- Stern, D. (2010). *Forms of Vitality. Exploring Dynamic Experience in Psychology, the Arts, Psychotherapy and Development*. Nueva York: Oxford University Press.
- Tomasello, M. (2005). Comments [to Everett, D. (2005). Cultural constraints on *pirāha* grammar. Another look at the design features of human language]. *Current Anthropology*, 46(4), pp. 621-646.
- Trevarthen, C. (1980). "The foundations of intersubjectivity: development of interpersonal and cooperative understanding in infants". En D. R. Olson (Ed.), *The Social Foundation of Language and Thought* (pp. 316-342). Nueva York: Norton.
- Van de Carr, F. R. y Lehrer, M. (1988). "Prenatal university: Commitment to fetal-family bonding and the strengthening of the family unit as an educational institution". *Pre- and Peri-natal Psychology Journal*, 3(2), pp. 87-102.
- Volterra, V., Caselli, M. C., Capirci, O. y Pizzuto, E. (2005). "Gesture and the emergence and development of language". En M. Tomasello y D. Slobin (Eds.). *Beyond Nature-Nurture. Essays in Honor of Elizabeth Bates* (pp. 3-40). Nueva Jersey: Lawrence Erlbaum.

- Vygotski, L. (1931/2000). "Historia del desarrollo de las funciones psíquicas superiores". En L. Vygotski, *Obras escogidas*, t.III (pp. 11-340). Madrid: Visor.
- Werner, H. y Kaplan, B. (1963). *Symbol Formation*. Hillsdale: Lawrence Erlbaum.

Desarrollo intersubjetivo y perceptivo

El ritmo como ejemplo de su enlace durante el primer año de vida

MAURICIO MARTÍNEZ (UAI)

1. Introducción

Podría pensarse que el desarrollo intersubjetivo y el desarrollo perceptivo se encuentran íntimamente vinculados, o bien que no se relacionan más allá de ciertas intuiciones próximas al sentido común. Aunque recientemente comenzó a resaltarse el importante papel que desempeña el desarrollo de la percepción intersensorial⁷² en relación con el desarrollo social, emocional, comunicativo y lingüístico (Bahrick y Lickliter, 2012), la naturaleza de esta relación aún no fue lo suficientemente explorada. General-

⁷² A lo largo del capítulo, respetaremos la terminología usada en el ámbito de estudios sobre el desarrollo de la percepción multisensorial (Bremner, Lewkowicz y Spence, 2012). Utilizaremos el término “percepción multisensorial” para referirnos a la capacidad para tratar con estímulos que son fuente de múltiple información sensorial. El término “intersensorial” refiere a la capacidad para percibir relaciones entre información que se encuentra disponible simultáneamente para diferentes sistemas perceptivos. Para describir la cantidad de información simultánea que ofrece un estímulo particular, utilizaremos los términos: “unimodal”, cuando presenta información disponible para un único sistema perceptivo; “bimodal”, para cuando presenta información para dos sistemas perceptivos; y “multimodal”, en aquellos casos en los que presenta información para tres o más sistemas perceptivos.

mente, los datos empíricos sobre las capacidades de percepción intersensorial de los bebés -obtenidos en estudios controlados de laboratorio- son utilizados como insumos en trabajos que abordan el desarrollo de distintos dominios: el aprendizaje de la relación entre palabras y objetos (Gogate, Bolzani y Betancourt, 2006), la emergencia de interacciones triádicas (Roszmanith, Costall, Reichelt, López y Reddy, 2014), el desarrollo de la intersubjetividad secundaria (Fogel y DeKoeyer-Laros, 2007), la emergencia del *sí mismo* (Stern, 1991) o la relación entre las interacciones sociales tempranas y el desarrollo social del bebé en edades más avanzadas (Jaffe, Beebe, Feldstein, Crown, Jasnow, Rochat y Stern, 2001). Asimismo, dichas capacidades también suelen ser invocadas en trabajos de índole teórica (Stevanovic y Koski, 2018) y filosófica (Verhage, 2008) relativos al contacto intersubjetivo. Si bien en los trabajos mencionados se destaca la relación entre la habilidad estudiada y la capacidad perceptiva, en ellos solo se da por sentada la relación sin explicitarse o conceptualizarse el vínculo.

Por otra parte, cuando se alude al desarrollo intersubjetivo, generalmente, el punto álgido de discusión suele ser el tránsito desde la intersubjetividad primaria -interacciones diádicas bebé-adulto- a la secundaria -interacciones triádicas bebé-adulto-objeto- (Trevarthen, 1998). En el ámbito de la psicología del desarrollo, se alude a este tránsito como la revolución del noveno mes (Rochat, 2004; Tomasello, 2013). La mayoría de las teorías que intentan explicar el desarrollo intersubjetivo temprano se ocupan, casi exclusivamente, del salto cualitativo desde la intersubjetividad primaria a la secundaria (M. Martínez, 2010). Sin embargo, el desarrollo intersubjetivo durante el primer año de vida se caracteriza, más que por saltos cualitativos o discontinuos, por cambios paulatinos y graduales en el modo con el que el bebé comparte su experiencia con

las personas (Reddy, 2008). En efecto, la participación del bebé durante el encuentro intersubjetivo evidencia otros cambios más allá del que ocurre durante la revolución del noveno mes. A lo largo del primer año de vida, los bebés desarrollan modos de contacto intersubjetivo con sus figuras de crianza cada vez más complejos y sutiles. Sus modos de participación cambian ostensiblemente durante la interacción.

Contamos con evidencia empírica suficientemente sólida que permite afirmar que, desde el nacimiento, los bebés presentan sesgos perceptivos que guían selectivamente su atención hacia aquellos parámetros de estimulación del medio que se sitúan precisamente en torno a las características físicas que definen perceptivamente a las personas (Rivière, 2003). Entre el mes y medio y los 6 meses, las conductas expresivas del bebé (sonrisas, vocalizaciones, movimientos), que inicialmente se manifiestan de manera breve y aislada, se van organizando progresivamente en forma de racimos (Kaye, 1986). Entre los 2 y 6 meses, los bebés pasan de solo atender visualmente a los gestos rítmicos ostensivos con objetos realizados por el adulto, a realizar ellos los primeros gestos canónicos rítmicos con objetos (Moreno-Núñez, Rodríguez y Del Olmo, 2015). A partir del segundo mes y hasta el octavo, expresan de diversos modos (conjugando vocalizaciones, expresiones faciales y gestos) su intención de volver a involucrar al adulto cuando este quiebra la dinámica de la interacción (Melinder, Forbes, Tronick, Fikke y Gredebäck, 2010). A partir de los 6 meses, los bebés se convierten en activos provocadores con la intención de atraer la atención del adulto hacia la interacción (Reddy, 2008). Entre los 6 y 12 meses, responden de manera diferenciada ante distintas expresiones de la conducta del adulto (imitación o entonamiento afectivo) durante la interacción (Bordoni,

Español y De Grande, 2017). Durante los últimos 3 meses del primer año de vida, los bebés comienzan a coordinar actos práxicos conjuntos y actos interpersonales (Hubley y Trevarthen, 1979), lo que da lugar a la emergencia de la atención conjunta (Tomasello, 2013) y la comunicación gestual preverbal (Sarriá y Rivière, 1991), ambos indicadores del establecimiento cabal de la experiencia de intersubjetividad secundaria.

Si asumimos que existe un vínculo entre la percepción intersensorial y el desarrollo intersubjetivo, resulta plausible interpretar que los cambios en la conducta del bebé se encuentran, de algún modo, sustentados por cambios en sus capacidades perceptivas. Ahora bien, ¿con qué clase de habilidades perceptivas se vinculan los cambios anteriormente mencionados?, ¿qué clase de cambios se producen en el desarrollo perceptivo vinculados a los cambios en el contacto intersubjetivo?, ¿qué clase de información percibe el bebé durante los intercambios? A su vez, la propia relación entre el desarrollo de ambas capacidades puede ser problematizada: debemos plantear una relación lineal-causal entre las capacidades perceptivas y el contacto intersubjetivo (o viceversa), o más bien debemos plantear una relación dialéctica (como sugieren Español, 2010 y Rivière, 2003), donde el progresivo involucramiento del bebé con el adulto propiciaría el desarrollo de la percepción intersensorial, y este impactaría en los cambios que se observan en el bebé en su involucramiento con sus figuras de crianza. Este es el tipo de preguntas que hay que responder si pretendemos establecer un vínculo que vaya más allá de dar por sentada la relación entre el desarrollo intersubjetivo y el desarrollo perceptivo.

Estas preguntas, además de obligarnos a profundizar nuestro conocimiento sobre el desarrollo de la percepción intersensorial, conducirían a la elaboración de

una explicación posible sobre el desarrollo intersubjetivo temprano, sorteando algunas dificultades. A grandes rasgos, podemos identificar dos tipos de problemas. En primer lugar, las teorías que pretenden explicar el tránsito de un modo de intersubjetividad a otro presentan, desde el punto de vista de su ontogénesis, algunos inconvenientes (M. Martínez, 2010). Por ejemplo: el innatismo explícito de Trevarthen, quien acuñó los conceptos de “intersubjetividad primaria” y “secundaria”; la concepción modular y maduracionista que subyace al modelo propuesto por Baron Cohen; o la disociación entre la capacidad general de inteligencia sensoriomotriz (en cuanto sustrato cognitivo) y las capacidades socio-comunicativas en la propuesta de Tomasello. En segundo lugar, también existen dificultades en las explicaciones acerca de cómo se relacionan los intercambios intersubjetivos tempranos con los denominados “precursores de la teoría de la mente”: los gestos deícticos y el juego de ficción (Baron Cohen, 1991).

Con relación al primer tipo de problemas, no contamos con una explicación apropiada sobre el desarrollo intersubjetivo que, además de indicar la existencia del tránsito de la intersubjetividad primaria a la secundaria, contemple los diferentes cambios que ocurren en la conducta interactiva del bebé durante el primer año de vida. Y con respecto al segundo problema, aunque algunos investigadores consideran el contacto intersubjetivo temprano como un precursor más de la teoría de la mente, anterior a los gestos deícticos, creemos que esto no es así. Bajo esta óptica, la interacción sería un eslabón más en la cadena genética que desemboca en *el proceso psicológico responsable de la comprensión de los demás*. Creemos que este modo de plantear el problema no lo resuelve. Contrariamente, asumimos que:

Las experiencias de intersubjetividad primaria no son un precursor de la teoría de la mente sino un conjunto de prácticas y capacidades corporeizadas que, si bien aparecen temprano en el desarrollo, continúan siendo la forma básica de entender a los otros en situaciones de interacción (Español, 2008, p. 130).

Por lo tanto, no deben entenderse como precursoras; por el contrario, esas experiencias deben explicarse ontogenéticamente en cuanto modos emocionales, sensoriomotrices, perceptuales y no conceptuales de compartir un tipo de experiencia que nos acompaña a lo largo de toda la vida (Gallagher, 2001).

El objetivo que nos proponemos en este escrito es presentar un posible enlace entre el desarrollo intersubjetivo y el desarrollo de las capacidades perceptivas de los bebés. Vincularemos el desarrollo de las capacidades de percepción intersensorial –estudiadas en contexto de laboratorio– con las conductas interactivas de los bebés –estudiadas en contextos observacionales y microanalíticos– para comprender la ontogénesis de los modos cada vez más sutiles y ajustados de interacción del bebé. En este trabajo en particular, nos focalizaremos en uno de los elementos constitutivos de las interacciones: el ritmo. Sabemos que el enlace también debería realizarse con otros elementos que conforman la interacción (la sincronía, la duración, la *rate* o densidad cronométrica), pero, por cuestiones de espacio, nos focalizaremos en el ritmo.

2. De los procesos que permiten inferir o simular a los procesos perceptivos

Es menester de la psicología identificar aquellos procesos que nos permiten comprender a los demás e involucrarnos socialmente con ellos, es decir, identificar aquellos

procesos responsables de la cognición social (Gómez y Núñez, 1998). Asimismo, es menester de la psicología del desarrollo dar cuenta de la ontogénesis de dichos procesos e identificar los cambios que se suceden en los diferentes modos de contacto social que se despliegan a lo largo del ciclo vital. En el año 1978, dos primatólogos, Premack y Woodruff, acuñaron una expresión que generó muchas hipótesis de trabajo y datos empíricos, lo que amplió nuestra comprensión sobre la cognición social. La expresión en cuestión es “*teoría de la mente*” (en adelante “TM”). Desde entonces, la TM se convirtió en el proceso psicológico que explica nuestra capacidad de interactuar, cooperar, engañar y tantas otras habilidades que ponemos en juego cuando nos vinculamos con los demás. En tal sentido, hoy día resulta difícil referirse a la cognición social o a cómo las personas interactúan eludiendo dicho término (Reddy y Morris, 2009). La TM es un proceso psicológico que nos permite interpretar, predecir y dotar de sentido a la conducta de los demás atribuyéndoles estados mentales, como creencias, deseos, etc. (Baron Cohen, Tager-Flusberg y Lombardo, 2013).

A partir de mediados de la década de 1980, se desarrollaron diferentes propuestas teóricas para explicar el funcionamiento y desarrollo de la TM. A principios de la década de 1990, la publicación de un número doble de la revista *Mind and Language* (1992, vol. 1-2) dio lugar a que se erigieran las dos grandes perspectivas que albergarían las diferentes propuestas teóricas, lo que le abrió paso a un debate que perdura hasta la fecha (Balmaceda, 2018). En el ámbito de la filosofía (Gallagher, 2001; Gomila, 2003; Pérez, 2013; Scotto, 2002) y en el de la psicología (Español, 2008), se denominan “perspectiva de primera persona” (PPP) y “perspectiva de tercera persona” (PTP), respectivamente. Para ambas, la interacción con las personas depende

de algún mecanismo psicológico que permita vislumbrar aquellos estados mentales -no accesibles a través de nuestros sistemas perceptivos- causantes de la conducta observable. Estas perspectivas propusieron sendas explicaciones sobre mecanismos que permiten inferir o simular los estados mentales a través de los cuales comprendemos la conducta de los demás. Por cuestiones de espacio, nos limitamos a nombrar algunas de las teorías más representativas del debate entre la PPP y la PTP. En M. Martínez (2011), el lector puede encontrar un análisis pormenorizado de un conjunto más amplio de propuestas teóricas.

Bajo la PTP encontramos propuestas como la de Leslie, quien planteó la existencia de un módulo responsable de metarrepresentaciones, representaciones mentales no literales en las que están suspendidas las relaciones normales de referencia y verdad respecto a los "estados del mundo" (Leslie, 1988). Bajo una óptica similar, Baron Cohen (2005) propuso un módulo con diferentes mecanismos que "maduran" a diferentes edades: detector de emociones, detector de intencionalidad, detector de dirección de la mirada, mecanismo de atención compartida, mecanismo de empatía y mecanismo de teoría de la mente. Por su parte, siguiendo la lógica de las metarrepresentaciones, Perner (1994) planteó el desarrollo de una capacidad intelectual general que permite operar con modelos que representan la relación representacional entre un modelo y el entorno. Una propuesta algo diferente es la de Wellman (2017), para quien la comprensión de los demás depende de un conocimiento teórico (inferencial), un conjunto de conceptos (deseos, creencias, etc.), distinciones ontológicas (entre lo real y lo pensado) y un marco causal-explicativo (el rol causal de los estados mentales en la conducta de las personas), sujetos, a su vez, a sucesivos cambios conceptuales. Bajo el amparo de la PPP, se encuentra

la propuesta de Harris (1991; 1992), quien atribuye al desarrollo de la imaginación la capacidad de proyectar los propios estados mentales a los demás para comprender cómo se comportarían en diferentes situaciones.

Mientras se desarrollaba el debate entre la PPP y la PTP, se intentaron algunas soluciones proponiendo modelos híbridos (Carruthers y Smith, 1996). A la par, emergió una segunda perspectiva (intentando elucidar algunos problemas no resueltos por estas dos perspectivas) denominada “perspectiva de segunda persona” (PSP) (Español, 2008; Gómez, 1996; Gomila, 2003; Pérez, 2013; Reddy, 1996; Scotto, 2002). Bajo la PSP se asume que experimentamos la mente del otro de manera directa e inmediata, sin necesidad de teorías, inferencias o proyecciones simuladas; percibimos la mente de los demás en sus expresiones en la medida en que nos involucramos en una interacción (Gomila y Pérez, 2017). Comprender a los demás desde la PPP y la PTP implica detectar (inferir o imaginar) que, por detrás de la conducta, en el plano de lo mental inobservable, hay fenómenos (creencias, deseos, etc.) que causan la conducta observada. En cambio, desde la PSP, comprender la conducta del otro implica percibir que existen en ella estructuras o formas (*Gestalten*) que la vinculan de manera directa con las otras personas (Gómez, 2008, 2009). Dicho vínculo no pertenece a un plano inobservable (como sucede con los deseos o las creencias); por el contrario, pertenece al mismo plano –el observable– que la conducta (Gómez, 2010). Gomila y Pérez (Gomila, 2003, Gomila y Pérez, 2017, Pérez y Gomila, 2018) postularon una serie de características que especifican la PSP:

1. La comprensión de las personas se produce solo cuando se vivencian interacciones dinámicas cara a cara o, mejor dicho, cuerpo a cuerpo. En tal sentido, es una comprensión *online*: la presencia física del otro es necesaria para su comprensión.
2. Los aspectos expresivos del cuerpo son vistos como expresivos (no interpretados), y por ello no puede ocurrir la comprensión si no es de cuerpo presente.
3. La comprensión es recíproca. Durante la interacción, la comprensión de la conducta por parte de uno de los participantes indefectiblemente modifica la conducta del otro.
4. El ejemplo paradigmático de este modo de comprensión no ocurre cuando entran en juego las actitudes proposicionales (deseos o creencias), sino cuando se trata de afectos dinámicos y emociones.
5. Las interacciones de segunda persona no suponen que haya una actividad *meta*, no hay un estado mental acerca del estado mental del otro, sino un estado mental causado por el estado mental-expresado-corporalmente del otro, y así sucesivamente.
6. Este proceso está ligado a la acción en el sentido de que involucra acción corporal pública, abierta, conducta emocional, *i.e.* alteraciones del rostro y otras conductas, como huida, consuelo, etc.
7. No es indispensable la existencia de un mundo de referencias compartidas.
8. No requiere del lenguaje.
9. El tipo de comprensión que se establece desde la PSP es anterior filogenéticamente, ontogenéticamente y lógicamente a la comprensión que se propone en la PPP y la PTP.

10. Este modo de comprensión, al ser anterior, resulta prioritario para la adquisición de los conceptos mentales involucrados en las atribuciones prototípicas de las PPT y PTP.

Al igual que en la PPP y la PTP, en la PSP se elaboraron algunas teorías que intentan explicar cómo se produce la comprensión o cognición social. Basándose en la fenomenología de Merleau-Ponty, Gallagher (2008, 2015) desarrolló la teoría de la interacción, la cual sostiene que la mayoría de los encuentros entre las personas depende de interacciones corporeizadas. Asume el concepto de “intercorporeidad” -propuesto por Merleau-Ponty-, según el cual las personas nos encontramos acopladas dinámicamente durante las interacciones intersubjetivas. La idea central es que en el encuentro con el otro coordinamos nuestras secuencias de percepción-acción; nuestros movimientos están acoplados con cambios en la velocidad, la dirección y la dinámica de los movimientos del otro. La idea de Gallagher que más nos interesa resaltar es que la percepción directa de las intenciones y emociones del otro contribuye de manera crucial a los modos de contacto intersubjetivo que Trevarthen (1998) describió como primarios y secundarios. Por su parte, Di Paolo y De Jaegher, a partir de una crítica desde el enactivismo (Varela, Thompson y Rosch, 1994) hacia la perspectiva cognitiva clásica, propusieron que la cognición social resulta de la coordinación interactiva que ocurre entre dos organismos, y no como consecuencia de una capacidad individual, como por ejemplo la TM (DiPaolo, Rohde y DeJaegher, 2010). Bajo su óptica, los bucles de percepción-acción que tienen lugar durante la interacción desempeñan un papel fundamental (Di Paolo, Rohde y Iizuka, 2007).

Como puede observarse, desde la PSP no se postulan mecanismos o procesos psicológicos modulares –como lo hace Leslie (1988)– o de dominio general –como lo hace Perner (1994)– que permitan realizar inferencias o simulaciones. Por el contrario, el proceso psicológico al que se apunta es a la percepción. Desde una perspectiva filosófica, y aludiendo a los enfoques interactivos de Gallagher, Di Paolo y De Jaegher, Scotto afirma que la comprensión de los demás “se desarrolla desde el nacimiento, depende fuertemente de la *percepción* y se implementa en contextos prácticos, es decir, se trata de una ‘capacidad pragmática’ que no presupone ni requiere ‘la capacidad para verbalizar las razones’” (2017, p. 94, la bastardilla nos pertenece). Por su parte, desde la perspectiva de la psicología del desarrollo, Español alude al contacto intersubjetivo desde la PSP diciendo: “Se trata de una expresión de cercanía, no basada en metarrepresentaciones de los otros sino en la *percepción* del otro como emocionalmente expresivo y dirigido al mundo” (2008, p. 127, la bastardilla nos pertenece).

Si, tal como reclamamos al inicio de este apartado, le corresponde a la psicología identificar los procesos que nos permiten comprender a los demás e involucrarnos socialmente, y si asumimos la PSP para dar cuenta de ello, la percepción se ubica en el centro de la escena. Asimismo, si la psicología del desarrollo debe explicar la ontogénesis de dichos procesos e identificar los cambios en la conducta del bebé que dan cuenta de su desarrollo intersubjetivo, una vez asumida la PSP se torna indispensable indagar el desarrollo de aquellos procesos perceptivos íntimamente involucrados en los intercambios sociales. Si bien existen algunos intentos de abordar al desarrollo intersubjetivo desde la PSP, como por ejemplo la hipótesis de una noción puramente sensomotriz del otro sugerida

por Gómez (2005), hasta donde sabemos, no conocemos propuestas que lo hayan vinculado de manera explícita con el desarrollo perceptivo.

3. Cuando la perspectiva de segunda persona aún no había sido esbozada, algunos observaron sabiamente la interacción

A partir de la década de 1960, comienzan a desarrollarse diferentes investigaciones que cambiaron la visión que se tenía hasta el momento de los bebés durante su primer año de vida. Por ejemplo, Fantz (1961) inaugura el estudio de las capacidades perceptivas tempranas con su estudio sobre el origen de la percepción de la forma. Más tarde, durante la década de 1970, se produce un cambio en el modo de concebir el desarrollo de los niños más pequeños (Rochat, 2004), en gran parte debido a una importante cantidad de datos empíricos surgidos de la investigación referida a la cognición social temprana (Case, 1989). Comienzan a realizarse estudios que analizan, por un lado, la conducta de los adultos frente a los bebés (ver Carretero y Español, 2016 para una revisión) y, por otro, la conducta interactiva que se despliega entre adultos y bebés (Schaffer, 1977).

Los trabajos que se focalizaron en el estudio de la conducta interactiva de los bebés y sus figuras de crianza anticiparon por casi más de veinte años las ideas que en la actualidad configuran lo que podríamos denominar el “núcleo duro de la PSP”. Los trabajos desarrollados por autores como Bruner, Bateson, Trevarthen, Stern, Kaye, Schaffer, Fogel, entre otros tantos, estudiaron la interacción tratando la conducta social en términos diádicos, enfatizando la dimensión temporal de las situaciones interactivas, a través de técnicas de microanálisis (Schaffer, 1977).

Mientras que, en los últimos años, DeJaegher y DiPaolo (2007) consideran la interacción como un nivel emergente y como un sistema autónomo transitorio (que dura mientras acaece la interacción), y mientras que Gallagher (2017) propone que la comprensión de los demás procede del intercambio intersubjetivo, Kaye proponía hace más de 30 años “la idea de que los padres y el bebé constituyen un ‘sistema social’ y la idea de la ‘intersubjetividad’ o comprensión compartida entre individuos” (1986, p. 47).

Con respecto al rol de la percepción, Schaffer sostuvo que, “en vez de buscar impulsos que ‘expliquen’ el desarrollo social, resulta más útil concebir que la sociabilidad se origina en los *encuentros perceptivos* del niño con otras personas” (1983, p. 68, la bastardilla nos pertenece). De igual modo, Hobson (1995) –uno de los primeros en sospechar de la necesidad de postular capacidades teórico-inferenciales o simulacionistas para la interacción social– supo poner el foco sobre la percepción. Su concepción acerca de cómo los bebés comprenden a los demás también anticipó temporalmente una de las críticas centrales esgrimida por la PSP para con la PPP y la PTP. Afirmó: “Si partimos de una escisión radical entre la ‘percepción del cuerpo’ y la ‘comprensión de la mente’, nunca llegaremos a juntar de nuevo las piezas de ese rompecabezas” (1995, p. 133). Propuso que los bebés perciben la cualidad de la actitud de la otra persona con respecto a su referente (el cual puede ser un objeto o el bebé mismo), sosteniendo que “muchas de las capacidades infantiles de relación social tienen algún tipo de fundamento perceptivo” (Hobson, 1995, p. 59), y que, “si ya a los 2 meses los bebés son capaces de participar en ciertas clases de intercambios interpersonales [...], deben contar con un sistema perceptivo suficientemente ajustado” (Hobson, 1995, p. 60).

Como puede observarse, desde las diferentes tradiciones dedicadas al estudio de las interacciones adulto-bebé también se alude a la percepción como mecanismo responsable del intercambio social temprano. En tal sentido, nos parece más que justificado embarcarse en la empresa de teorizar, de manera sistemática y explícita, las relaciones entre el desarrollo intersubjetivo temprano y el desarrollo de la percepción.

4. La conducta del adulto durante el encuentro intersubjetivo

En los apartados anteriores, dejamos en claro que la percepción parece desempeñar un rol crucial tanto para el encuentro intersubjetivo, como para el propio desarrollo de la intersubjetividad. No obstante, una cuestión que resulta esencial para comprender el vínculo entre el desarrollo de la percepción y el desarrollo intersubjetivo temprano es la concerniente a qué percibe el bebé durante el encuentro intersubjetivo. Por ejemplo, desde la PSP se sostiene que, durante los intercambios intersubjetivos, los participantes perciben de manera directa las emociones y las intenciones de los compañeros de interacción. Ahora bien, si nos detenemos a observar la interacción entre un adulto y un bebé, resulta que el concepto de “emoción”, por ejemplo, no es buen ejemplo para caracterizar lo que el bebé percibe. Tradicionalmente, este concepto se suele asociar directamente con la manifestación de un conjunto de expresiones faciales: ira, felicidad, tristeza, asco, sorpresa y miedo (Ekman, 1982). Sin embargo, Stern (1983) describió cómo los adultos, cuando se encuentran frente a los bebés, expresan las emociones en su rostro de un modo distinto al que lo hacen cuando se encuentran frente a otro adulto. Frente al bebé, el adulto exagera la configuración

de los elementos que componen la expresión facial, por ejemplo, los ojos y la boca se abren con mayor amplitud, mientras que la conformación de la expresión facial se caracteriza por su lenta formación y su prolongada duración. En tal sentido, durante la interacción, desde la perspectiva del bebé, resulta más llamativo el devenir temporal de la configuración de la expresión facial –su lento florecimiento y su paulatino desvanecimiento– que la configuración en sí (Stern, 1991).

Si bien las emociones discretas, como así también el habla dirigida al bebé, tuvieron su apogeo como objeto de estudio y análisis, el devenir histórico del estudio de la actuación adulta recaló en la noción de *performance* artística dirigida al bebé (Carretero y Español, 2016), la que Español y Shifres (2015) proponen como unidad de análisis apropiada para caracterizar y comprender la conducta adulta frente a bebés.

Para atraer y sostener la atención del bebé hacia la actividad de interacción, el adulto ofrece información sensorial para varios de los sistemas perceptivos de aquél: visual, auditivo y táctil (Stern, Beebe, Jaffe y Bennett, 1977). Los adultos le hablan al bebé, lo tocan, lo mueven y se mueven delante de él. De modo similar a como ocurre con las expresiones faciales, la información que provee el adulto presenta también, en este caso, ciertos rasgos particulares: la lentitud de los cambios y una organización regular en el tiempo (Stern, 1983). Por ejemplo, el habla del adulto presenta rasgos melódicos prolongados en el tiempo con alturas más graves y más agudas que las del habla que un adulto dirige a otro adulto (Papoušek y Papoušek, 1981). Si bien la conducta adulta puede contener habla, chistidos o vocalizaciones (o bien prescindir de toda información audible), lo importante es destacar que dicha información sonora, cuando está presente, suele

acompañarse de toques y movimientos coordinados sincrónicamente (I. Martínez, Español y Pérez, 2018; Ospina y Español, 2014).

La conducta que el adulto ofrece al bebé frecuentemente se organiza en episodios de participación mutua compuestos por *frases* (emisiones únicas de voz y/o ráfagas de movimientos), *series* (varias frases de contenido o duración semejantes) y episodios de *pausa* que consisten en un silencio comportamental sonoro y kinético (Stern *et al.*, 1977). Las series se despliegan en el tiempo según la forma repetición-variación (Español, 2014), brindando al bebé una estimulación lo suficientemente iterativa, estable y relacionada temporalmente con su propia conducta como para que este pueda percibir con claridad la relación de contingencia que evocan su propias respuestas sociales, pero también lo suficientemente variable como para sostener la atención del bebé durante la interacción (Rivière, 2003).

Desde el inicio, la información que provee el adulto conforma patrones multimodales: los contornos melódicos están estrechamente relacionados con los *patterns* de movimiento de la madre, y la sincronización regular de *patterns* vocálicos y kinéticos -que incluyen información táctil (cuando lo acaricia, lo toca o lo empuja), propioceptiva (cuando le mueve las manos o los pies) y visual (al saludarlo con la cabeza o cuando agita la cabeza)- provee al bebé de información redundante para todos sus sistemas perceptivos (M. Papoušek, 1996).

La observación de la conducta adulta a través de herramientas teóricas y técnicas provenientes de las artes temporales (música, danza y poesía) resaltaron la importancia de su moldeado dinámico, temporal y multimodal (Schögler y Trevarthen, 2007; Miall y Dissanayake, 2003; I. Martínez, 2014; Shifres, 2014; Español y Shifres, 2015;

I. Martínez, Español y Pérez, 2018). Estos estudios destacaron el rol fundamental que desempeña la información temporal (duración, *rate*, ritmo e intensidad) para la interacción. El adulto elabora intuitivamente los sonidos y los movimientos que realiza (H. Papoušek, 1996). Los adultos *manipulan* el ritmo (Koester, Papoušek y Papoušek, 1989), la duración (Stern, 1983) e intensidad (Stern, Hofer, Haft y Dore, 1985) con el fin de crear una *performance* que capture y sostenga la atención del bebé hacia el contacto intersubjetivo. Durante la interacción, el bebé participa en una *performance* multimodal modelada temporalmente en función del ritmo, la duración y/o la intensidad de la información sonora, visual, etc.

Durante la interacción, el bebé percibe dos tipos de información, la información de modalidad única y la información amodal. La información de modalidad única es aquella que solo puede ser percibida por un único sistema perceptivo, mientras que la información amodal es aquella que se percibe a través de diferentes sistemas perceptivos (Bahrick y Lickliter, 2012). Describamos una situación prototípica de interacción para ejemplificar cada tipo de información. Una niña de 7 meses se encuentra recostada delante de su papá. El papá toca suavemente y de manera pulsátil con su dedo índice la nariz de su hija al tiempo que vocaliza el sonido /tic/. Los breves y suaves toques y las vocalizaciones del padre se producen reiteradamente de manera simultánea y sincrónica ajustados a un patrón rítmico. Los gestos faciales, el color del pelo, el timbre de voz, los toques del dedo, el aroma del perfume y las vocalizaciones del padre es información que solo puede ser percibida por un único sistema perceptivo. Los gestos faciales y el color del pelo solo se perciben a través de la visión. El timbre de voz y las vocalizaciones solo se perciben a través de la audición. El toque del dedo solo se percibe

por el tacto pasivo. El perfume solo se percibe a través del olfato. Ahora bien, ¿qué ocurre con el ritmo? El /tic/ que vocaliza el padre y el toque en la nariz del dedo ocurren de manera sincrónica y con el mismo patrón rítmico. El ritmo de la vocalización y del toque se percibe indistintamente a través de la audición y del tacto. Más aún, si la niña focaliza en ese momento su atención visual en el dedo, el ritmo también puede percibirlo visualmente en el movimiento del dedo.

De esta manera, durante su participación en el intercambio intersubjetivo, el bebé percibe dos tipos de información. En tal sentido, si asumimos las ideas propuestas por Schaffer o Hobson –respecto a que la sociabilidad se origina en los encuentros perceptivos del bebé con otras personas, o que muchas de las capacidades infantiles de relación social tienen algún tipo de fundamento perceptivo–, el desarrollo de la capacidad para detectar o discriminar estos dos tipos de información, como así también el desarrollo de la habilidad para reconocer relaciones entre ellas, resulta una cuestión crucial para comprender el desarrollo intersubjetivo temprano. El estudio del desarrollo perceptivo sería así una vía válida para intentar comprender los cambios en la participación del bebé durante la interacción a lo largo del primer año, prescindiendo de la necesidad de aludir a procesos psicológicos responsables de inferencias o simulaciones que permitan hacer visibles los opacos estados mentales causantes de la conducta.

5. El desarrollo de la percepción intersensorial

En los dos apartados anteriores, destacamos la importancia atribuida a la percepción, al tiempo que mostramos la falta de una articulación teórica explícita entre su desarro-

llo y el desarrollo intersubjetivo temprano. Más aún, destacamos el carácter constitutivo de la multimodalidad para la *performance* adulta, resaltando, además, su organización temporal en términos de la redundancia intersensorial que brinda a través de la información amodal temporal: la sincronía, la duración, la *rate* y/o el ritmo. En tal sentido, resulta crucial comprender cómo interactúa el desarrollo de la percepción intersensorial con la actuación del adulto en cuanto *performance* multimodal.

La percepción intersensorial consiste en la capacidad de percibir de manera unificada y coherente la información que una única fuente estimular brinda, de manera simultánea, para dos o más sistemas perceptivos. El desarrollo de la percepción intersensorial enfrenta al bebé con un importante desafío evolutivo: ser capaz de atender de la forma más eficiente posible a la información multimodal relevante para sus necesidades y acciones, mientras ignora la estimulación menos relevante (Bahrick y Lickliter, 2012).

Los estudios de percepción intersensorial indagan el desarrollo de diferentes capacidades perceptivas: la transferencia intersensorial (Bahrick, 1988), la discriminación y el reconocimiento (Lewkowicz, 2000; Walker-Andrews, 1997). La transferencia intersensorial es la habilidad de identificar la misma información amodal presentada a través de diferentes sistemas perceptivos. Por ejemplo, identificar el mismo patrón rítmico presentado primero a través del sistema perceptivo visual y, luego, a través del sistema perceptivo auditivo. La discriminación es la habilidad para diferenciar dos eventos estimulares. Por ejemplo, identificar que dos patrones rítmicos son diferentes. El reconocimiento, por su parte, es la habilidad que permite establecer relaciones entre la información percibida, simultáneamente, a través de dos o más sistemas perceptivos. Por ejemplo, identificar que el patrón rítmico percibido a través del

sistema perceptivo visual es el mismo que el percibido a través del sistema perceptivo auditivo. Cada una de estas capacidades, a su vez, se estudia con diferentes técnicas de investigación (Bahrick, 1988; M. Martínez, 2016).

6. La hipótesis de la redundancia intersensorial

Como ya mencionamos, una de las características distintivas de la actuación adulta es su multimodalidad. En el ejemplo del padre con su beba y los toques rítmicos que realiza con su dedo sobre la nariz de su hija, se puede identificar otra característica esencial de la *performance* que emerge solo en función de su multisensorialidad: la redundancia intersensorial. La pequeña recibe de manera simultánea y sincrónica la misma información -el mismo patrón rítmico- para dos o más sistemas perceptivos; escucha el patrón rítmico en las vocalizaciones, observa el patrón rítmico en el movimiento del dedo de su papá y siente el patrón rítmico a través del tacto sobre su nariz. A este fenómeno se lo denomina “redundancia intersensorial”. La redundancia intersensorial se encuentra en un evento cuando presenta la misma información amodal (sincronía, duración, *rate*, ritmo) de manera simultánea y sincronizada temporalmente disponible para dos o más sistemas perceptivos (Bahrick y Lickliter, 2012). Se trata, por ende, de una propiedad de la estructura de los objetos y eventos multimodales del entorno (Bahrick y Lickliter 2002).

Respecto al rol que desempeña la redundancia intersensorial en el desarrollo de la percepción intersensorial, Bahrick y Lickliter (2012) plantearon la hipótesis de la redundancia intersensorial (HRI). La HRI establece que la atención selectiva guía el desarrollo perceptivo temprano. Proporciona un marco para comprender cómo y

en qué condiciones los organismos atienden a la información amodal en comparación con la información de modalidad única. La HRI da cuenta de cómo los bebés, sin conocimiento previo del mundo, perciben de manera unitaria y coherente los eventos atendiendo a la información que es relevante para sus necesidades y acciones. Bahrick (2004) plantea que la redundancia intersensorial incide en el organismo a nivel de los procesos atencionales. Los efectos de la redundancia tienen un alto impacto sobre la percepción, el aprendizaje y la memoria, lo cual deriva en formas diferenciadas de atención a las distintas propiedades de los estímulos. La hipótesis provee cuatro predicciones, dos referidas a la atención selectiva hacia diferentes propiedades (tipos de información) de los eventos y otras dos referidas a cambios evolutivos.

- *Predicción 1:* La información amodal redundante de un evento sobresale por encima de las demás propiedades y se detecta más fácilmente cuando se encuentra disponible de manera bimodal -para dos sistemas perceptivos: visual y auditivo, por ejemplo- que cuando se encuentra disponible para un único sistema perceptivo o de manera unimodal.
- *Predicción 2:* La información de modalidad específica de un evento sobresale por encima de las demás propiedades y se detecta más fácilmente en eventos que proveen información para un único sistema perceptivo, es decir, un evento unimodal.
- *Predicción 3:* Durante el desarrollo, se incrementa en el bebé su capacidad para la diferenciación perceptiva, lo que significa que la eficiencia del procesamiento y la flexibilidad atencional permite la detección de información amodal y de modalidad única en eventos unimodales y bimodales.

- *Predicción 4:* Tanto la facilitación intersensorial (predicción 1) como la facilitación unimodal (predicción 2) son más pronunciadas para aquellas tareas de alta dificultad en relación con la experiencia del organismo que percibe. Por lo tanto, se evidencia a lo largo de todo el ciclo vital.

La HRI tiene una serie de implicancias más que sugerentes para comprender el desarrollo intersubjetivo temprano. Una esencial es que las *performances* multimodales ofrecidas por el adulto, al presentar redundancia intersensorial, ponen en primer plano la información amodal (o propiedades amodales) y en segundo plano la información de modalidad única o propiedades unimodales. La redundancia intersensorial atrae y focaliza la atención del bebé hacia los eventos sociales y sus propiedades amodales (sincronía, ritmo, duración, intensidad), promoviendo su percepción de manera coherente y sirviendo de control para un mayor procesamiento perceptivo de eventos sociales cohesivos (Bahrick, 2010). Por ejemplo, durante el primer año de vida, los bebés se tornan cada vez más capaces de coordinar la mirada, sus vocalizaciones y sus movimientos durante la interacción social (Kaye, 1986). En tal sentido, Bahrick (2010) propone que la detección de la redundancia intersensorial de la información amodal (información que los adultos intuitivamente manipulan) durante los intercambios sociales juega un papel importante en el análisis de la corriente de estimulación, iniciando y manteniendo la atención infantil hacia la interacción social. Participar en interacciones sociales recíprocas depende de lograr la coordinación temporal entre los propios comportamientos y los del compañero de interacción (DiPaolo, Rohde y DeJaegher, 2010). La detección de información amodal y la redundancia intersensorial serían la base de este proceso

que los bebés desarrollan de manera muy rápida durante los primeros 6 meses de vida. Por ejemplo, en los primeros meses de vida, los bebés aprenden a participar en intercambios recíprocos e interacciones de turnos con sus cuidadores adultos, sus movimientos y ritmos vocálicos están intercoordinados con el patrón temporal de la comunicación adulta, y esto se basa en la percepción intermodal de relaciones propioceptivas-visuales-auditivas (Bahrick, 2010).

A pesar de la gran importancia de la HRI, a nuestro entender solo permite explicar y comprender por qué la *performance* del adulto atrae la atención del bebé, al tiempo que facilita la discriminación entre diferentes tipos de información amodal. Aunque la HRI permite comprender ciertos cambios en la conducta interactiva del bebé en función de su desarrollo perceptivo, por ejemplo, a partir de la pregnancia atencional de la *performance* en detrimento de eventos “no sociales” o la capacidad de discriminar información amodal en estímulos bimodales y unimodales a diferentes edades (Bahrick y Lickliter, 2012), estos no son los únicos cambios que tienen lugar durante el desarrollo perceptivo del bebé a lo largo del primer año de vida. Según hemos mencionado antes, la percepción intersensorial implica, al menos, otras dos capacidades perceptivas además de la discriminación: los bebés también desarrollan la capacidad de transferencia intersensorial y de reconocimiento. Sumado a esto, los adultos también suelen realizar *performances* en las cuales solo brindan información para una única modalidad sensorial (más adelante analizaremos un ejemplo de *performance* con esta característica).

7. El desarrollo de la percepción de información temporal

Lewkowicz (1994; 2000) propuso que la comprensión de relaciones intersensoriales basadas en distintos tipos de información amodal temporal emerge progresivamente durante el primer año de vida. Justamente, recordemos, es este tipo de información (sincronía, duración, *rate*, ritmo) la que los adultos manipulan con el fin de atraer y sostener la atención del bebé durante los intercambios intersubjetivos. Según el modelo, los bebés comprenden desde muy temprano (2 meses) la ocurrencia sincrónica de la información disponible para dos sistemas perceptivos, luego comprenden las relaciones basadas en la duración, a continuación emerge la comprensión de las relaciones basadas en la *rate*, y finalmente comprenden las relaciones basadas en el ritmo. Si bien la propuesta es sumamente interesante para comprender los cambios en el desarrollo intersubjetivo, creemos que presenta algunas limitaciones.

En primer lugar, Lewkowicz sostiene que los bebés son sensibles a diferentes relaciones intersensoriales, sin explicitar de qué tipo de sensibilidad está hablando. Es decir, no explicita a qué habilidades perceptivas (transferencia intersensorial, discriminación, o reconocimiento) se refiere. En segundo lugar, sostiene que a diferentes edades los bebés son sensibles a relaciones intersensoriales basadas en diferentes tipos de información, pero no dice qué ocurre con esa habilidad en edades posteriores. Afirma que entre los 3 y 6 meses los bebés son sensibles a las relaciones basadas en la duración, pero no indica nada respecto de dicha sensibilidad a los 10 meses. Esta cuestión es relevante porque el desarrollo de la percepción intersensorial no implica solamente la emergencia de diferentes habilidades respecto de diferentes tipos de información a diferentes edades. Muy por el contrario, el

desarrollo de la percepción intersensorial implica necesariamente un incremento en la efectividad de la habilidad (Bahrick, 2010; Gibson, 1969).

Intentando sortear la primera dificultad que mencionamos, en otro lugar (Martínez, Español e Igoa, 2018) propusimos una secuencia evolutiva relativa a la sensibilidad a las relaciones intersensoriales basadas en el ritmo. En el apartado que sigue, expondremos detalladamente esta secuencia. De igual modo, respecto a la segunda dificultad mencionada, hemos encontrado evidencia empírica que da cuenta de que los bebés de 10 meses son más eficientes que los de 4 meses para reconocer relaciones intersensoriales basadas en la duración (Martínez, Español e Igoa, en preparación).

Iniciamos este trabajo resaltando la necesidad de dar cuenta de una serie de cambios en la conducta interactiva del bebé que refleja su desarrollo intersubjetivo más allá del clásico tránsito desde la intersubjetividad primaria hasta la secundaria. Expusimos la importancia que se le concede a la percepción con relación al desarrollo intersubjetivo, tanto desde la PSP como desde los estudios iniciales de interacción adulto-bebé. Detallamos también la relevancia que tiene la información amodal temporal (sincronía, duración, *rate*, ritmo) para la interacción en cuanto la *performance* se conforma con esta información, a la par que es esta misma información la que es manipulada por el adulto para atraer y sostener la atención del bebé hacia la interacción. Finalmente, expusimos el rol que la redundancia intersensorial –una de las características constitutivas de la *performance* adulta– desempeña para explicar la orientación preferencial hacia las personas, así como la deriva evolutiva de diferentes habilidades de percepción intersensorial de información amodal temporal. En el apartado que sigue, daremos un paso más. Enlazaremos,

puntualmente, el desarrollo de distintas habilidades de percepción intersensorial del ritmo con algunos cambios descritos en la conducta interactiva del bebé a diferentes edades. Para ello nos valdremos de datos observacionales tomados de diferentes estudios de microanálisis.

8. El contacto intersubjetivo y la percepción del ritmo

Antes de adentrarnos en el enlace entre el desarrollo de la percepción intersensorial y los cambios de conducta del bebé a lo largo de la interacción a diferentes edades, durante el primer año de vida, analizaremos -en términos generales- la función del ritmo en la interacción social y profundizaremos un poco más en el desarrollo de diferentes habilidades de percepción intersensorial del ritmo.

El ritmo es considerado un elemento clave en el acople, enganche o arrastre (*entrainment*) que se observa en múltiples actividades en las cuales las personas comparten su experiencia (Phillips-Silver, Aktipis y Bryant, 2010). Puede observarse en la práctica de los deportes, en el juego, en la comunicación verbal, en el intercambio emocional, en la música y en la danza (McNeill, 1995). El ritmo sincroniza los procesos afectivos, perceptivos y motores de los sistemas nerviosos de los participantes de una interacción (Perinat, 1993). Trevarthen (1984) propuso que la empatía mutuamente consciente que se despliega entre los participantes (adulto y bebé) de un encuentro intersubjetivo temprano depende del ritmo que impregna el movimiento expresivo del compañero de interacción.

Desde antes del nacimiento y durante el primer año de vida, se despliega una serie de diferentes capacidades perceptivas. Existen datos empíricos que dan cuenta de que, hacia el final de la gestación, los fetos detectan y

discriminan patrones rítmicos en función de la información somatosensorial-táctil-vestibular, originada a través de la frecuencia respiratoria materna y de movimientos rítmicos antero-posteriores y laterales de la madre (Provasi, Anderson y Barbu-Roth, 2014). Desde el nacimiento, los bebés son sensibles (en cuanto detectan variaciones) a la fluctuación de patrones rítmicos, aunque no es posible determinar si son capaces de discriminar entre dichos patrones en función de su estructura jerárquica (Lewkowicz, 2000). Entre los 4 y 10 meses, los bebés discriminan entre dos patrones rítmicos presentados de manera bimodal (sonido y movimiento) (Pickens y Bharick, 1995; Lewkowicz y Marcovitch, 2006). A los 7 meses, pueden identificar el mismo patrón rítmico cuando primero se presenta en la modalidad visual y luego en la modalidad auditiva (Allen, Walker, Symonds y Marcell, 1977). A partir de los 8 meses, además de discriminar patrones rítmicos presentados de manera bimodal (visual y auditivo), los bebés también discriminan entre dos patrones rítmicos presentados en formato unimodal (visual) (Bahrick y Lickliter, 2004). Desde los 10 meses, los bebés discriminan entre dos patrones rítmicos bimodales aun cuando la información visual y auditiva se presenta de manera simultánea pero no sincrónica (Lewkowicz, 2003). A esa misma edad, los bebés reconocen relaciones intersensoriales basadas en el ritmo (Martínez, Español e Igoa, 2018).

Estos datos, obtenidos en estudios de laboratorio, dan cuenta del desarrollo de diferentes capacidades de percepción intersensorial vinculadas al ritmo. En virtud de ellos, resulta factible pensar que, a distintas edades, la conducta interactiva del bebé será diferente cuando el devenir temporal de la *performance* multimodal adulta se encuentre modelada rítmicamente. A continuación (ver Tabla 1), presentamos un inventario de diferentes comportamien-

tos descritos en estudios observacionales de interacciones, que dan cuenta de la conducta del bebé frente a una actuación adulta modelada rítmicamente. Las filas de la tabla 1 distinguen las interacciones con y sin mediación de objetos y el modo de participar del bebé y del adulto. Las columnas de las tablas identifican las edades en las cuales fueron realizadas dichas observaciones. En las cuadrículas que se forman, se realizan breves descripciones de las actuaciones de cada participante en función del modelado rítmico de la actuación adulta.

Tabla 1. La conducta del bebé y del adulto en relación con el desarrollo de la percepción intersetorial del ritmo

		2 meses	3 meses	4 meses
Interacción sin objetos	Adulto	Coordina su ritmo cardíaco con el del bebé durante la interacción (4).	Ofrece episodios ritmicos durante el 48,30 % de tres minutos de interacción (2).	Toma a la bebé por el torso y la sacude ritmicamente hacia arriba y hacia abajo (7).
	Bebé	Coordina su cardíaco con el de la mamá durante la interacción (4).	Orienta su atención visual hacia la conducta rítmica del adulto (2).	Es capaz de reproducir el pulso y las cualidades rítmicas de la <i>performance</i> adulta durante las pausas que éste realiza (7).
Habilidad perceptiva del bebé		Sensibilidad (detección) al ritmo.	Sensibilidad (detección) al ritmo.	Discriminación de patrones ritmicos bimodales.
	Bebé	Presta atención a la acción rítmica cuando se pone el objeto entre la línea del contacto visual (1).	Puede así el objeto, esto facilita la entrada en su uso. Realiza los primeros gestos ostensivos dirigidos a si mismo (1).	Continúa la preferencia por el uso de objetos sonoros. Usa los objetos ritmicamente. Realiza demostraciones distantes de los objetos. Realiza gestos ostensivos con y sin ritmo (1).
Interacción con objetos		Usa objetos sonoros y no sonoros. Usa más los sonoros. Ambos objetos son usados ritmicamente. Realiza demostraciones distantes de los objetos. Realiza gestos ostensivos con (1).		
	Adulto			

		5 meses	6 meses	7 meses
Interacción sin objetos	Adulto	Realiza <i>performance</i> con la forma repetición-varación de un patrón rítmico multimodal (3).		
	Bebé	Vocaliza conjuntamente con la madre. Cae en el pulso subyacente (3).		
Habilidad perceptiva del bebé		Discriminación de patrones rítmicos bimodales.	Discriminación de patrones rítmicos bimodales.	Discriminación de patrones rítmicos bimodales. Identificación del mismo patrón rítmico en dos modalidades sensoriales diferentes.
Interacción con objetos	Bebé	Primeros usos sonoros convencionales de los objetos sonoros. Aumenta la frecuencia de gestos ostensivos dirigidos a sí mismo. Frecuencia máxima de mirada del bebé hacia demostraciones distantes del uso sonoro de objetos (1). Aún no puede coordinar su atención visual con la de la madre (9).	Continúa con el uso rítmico y sonoro de los objetos. Decrece la frecuencia de demostraciones y gestos ostensivos (1).	Realiza un discurso hablado con una gran variedad rítmica que presenta rasgos musicales expresivos (9).
	Adulto			

		8 meses	9 meses	10 meses
Interacción sin objetos	Adulto	Inicia una acción conjunta lúdica que consiste en golpetear ritmicamente la mesa de la silla alta del bebé (6)	Ofrece una <i>performance unimodal</i> en medio de una gran <i>performance multisensorial</i> (5)	Continúa con el juego del golpeteo (6)
	Bebé	Parece no tener conciencia sobre su propio golpeteo (6)	Segmenta y anticipa la conducta multimodal y unimodal del adulto (5)	Durante el juego de golpeteo emerge la alternancia de la mirada (6)
Habilidad perceptiva del bebé	Adulto	Discriminación de patrones rítmicos bimodales. Identificación del mismo patrón rítmico en dos modalidades sensoriales diferentes.	Discriminación de patrones rítmicos bimodales. Identificación del mismo patrón rítmico en dos modalidades sensoriales diferentes.	Discriminación de patrones rítmicos bimodales. Identificación del mismo patrón rítmico en dos modalidades sensoriales diferentes.
	Bebé	Discriminación de patrones rítmicos unimodales.	Discriminación de patrones rítmicos unimodales.	Discriminación de patrones rítmicos unimodales. Reconocimiento de relaciones intersensoriales basadas en el ritmo.
Interacción con objetos	Bebé			
	Adulto			

Referencias: (1) Moreno-Núñez, Rodríguez y Del Olmo (2015), (2) Koester, Papoušek y Papoušek (1989), (3) I. Martínez (2007), (4) Feldman et al. (2011), (5) Espaniol (2004), (6) Fogel y DeKoeyter-Laros (2007), (7) Ospina y Espaniol (2014), (8) Rossmannith, Costall, Reichelt, López y Reddy (2014), (9) I. Martínez (2008).

Las tablas solo nos permiten contrastar algunas conductas interactivas del bebé (vinculadas con la actuación adulta modelada rítmicamente) y el desarrollo de diferentes capacidades de percepción intersensorial coevolutivamente consideradas. Pese a ello, resulta posible identificar en ellas algunas regularidades en el cambio de la actuación interactiva de los bebés frente a la conducta adulta modelada rítmicamente, que pueden enlazarse con el desarrollo de la percepción intersensorial del ritmo. En la tabla 2, esquematizamos dichas regularidades para las interacciones sin objeto, y, en la tabla 3, esquematizamos las regularidades para las interacciones con objetos. A diferencia de la tabla 1, aquí la información se circunscribe al bebé.

Tabla 2. Coevolución del desarrollo intersubjetivo en formato diádico y desarrollo perceptivo

<i>La performance</i> rítmica regula el estado emocional del bebé y nivel de activación.	La atención del bebé es atraída hacia la <i>performance</i> rítmica.	Discretiza la conducta adulta y la anticipa en función de <i>performances</i> rítmicas multimodales.	Discretiza la conducta adulta y la anticipa en función de <i>performances</i> rítmicas unimodales.	Introduce patrones rítmicos regulares en la interacción.	Se apoya en la <i>performance</i> rítmica para la alternancia visual.
<i>Sensibilidad al ritmo → Discriminación bimodal → Discriminación unimodal → Reconocimiento de relaciones intersensoriales.</i>					

Tabla 3. Coevolución del desarrollo intersubjetivo en formato triádico y desarrollo perceptivo

Su atención es atraída hacia la acción rítmica con el objeto.	Su atención sigue la secuencia de una acción rítmica con un objeto.	Realiza uso sonoro de los objetos.
<i>Sensibilidad al ritmo → Discriminación bimodal → Discriminación unimodal → Reconocimiento de relaciones intersensoriales.</i>		

Estos datos permiten formular nuevas preguntas para la realización de estudios observacionales tendientes a comprender el desarrollo intersubjetivo temprano en situaciones diádicas de juego social y en situaciones interactivas mediadas por objetos.

Los apartados siguientes están dedicados a avanzar un poco más en el enlace entre el desarrollo de la percepción intersensorial del ritmo y el desarrollo intersubjetivo temprano, explicitando aún más la relación entre la conducta interactiva del bebé y sus capacidades perceptivas.

8.1. El ritmo y la interacción a los 3 meses

Field (1978) propuso que los adultos modulan la conducta del bebé acompañando el ritmo de sus acciones y movimientos. Asimismo, Papoušek y Papoušek (1981) sostuvieron que los adultos despliegan una variedad de patrones rítmicos para modular o regular el estado conductual o emocional del bebé. En un estudio donde se analizaron los tres primeros minutos de una sesión de juego social entre bebés de 3 meses y sus mamás, se identificó que los movimientos rítmicos constituyen la mayor proporción de la conducta de los adultos durante la interacción (Koester, Papoušek y Papoušek, 1989). Si durante los primeros meses de su vida y hasta el cuarto mes, aproximadamente, los bebés demuestran sensibilidad al ritmo (Lewkowicz, 2000), se puede interpretar que la respuesta de los bebés frente a la estimulación rítmica que los adultos despliegan para modular o regular su conducta y sus emociones depende en gran medida de dicha sensibilidad.

8.2. El ritmo y la interacción entre los 7 y 10 meses

Según hemos dicho, hasta los 8 meses los bebés solo discriminan entre dos patrones rítmicos cuando estos son presentados a través de información visual y auditiva, es decir, bimodalmente. A partir de los 8 meses, desarrollan la capacidad para discriminar patrones rítmicos presentados de manera unimodal, es decir, visual o auditivamente. La capacidad para discriminar diferentes patrones rítmicos es posible que se encuentre íntimamente vinculada con el desarrollo de la capacidad para segmentar el flujo perceptivo que implica la conducta del otro durante la interacción. Según Español (2004), la reiteración de un secuencia rítmica permite al bebé descomponer (*discretizar*) la acción del otro y anticipar lo que vendrá. Ahora bien, ¿por qué resulta fundamental para esta habilidad la capacidad para discriminar entre patrones rítmicos diferentes? Sencillamente, creemos que es fundamental en virtud de que la forma repetición-variación es constitutiva de la *performance* adulta (Español, 2014). Recordemos que, bajo este formato, el adulto brinda al bebé una estimulación lo suficientemente repetitiva y estable para que pueda percibirla con claridad, pero también lo suficientemente variable como para sostener su atención. Si pensamos en una *performance* prototípica en la cual el motivo que se repite y varía es justamente un patrón rítmico, resulta claro el rol decisivo que juega la capacidad de discriminar entre diferentes patrones rítmicos al momento de discretizar la conducta del adulto.

Para mostrar cómo se enlaza el desarrollo de la percepción intersensorial con el cambio en la conducta interactiva del bebé, vamos a transcribir dos observaciones realizadas por Español (2004) en las cuales se observa cómo un niño de 9 meses segmenta y anticipa la conducta de

la adulta con la que se encuentra interactuando cuando ella realiza dos *performances* diferentes, una multimodal y la otra unimodal.

Observación 1: Cuando H tiene 0; 9 (02), inicio una de esas interacciones típicas que consta de cantar a la vez que se mueven las manos. H. está atento todo el tiempo, mirando mis manos o estableciendo contacto ocular. Cuando termino, H. me sigue mirando, y entonces yo repito el juego. Y así varias veces. Al rato, me acerco a él, poco a poco, produciendo sonidos rítmicos hasta golpear mi cabeza con él. H. me mira interesado y la escena transcurre con predominio de contacto ocular. De a poco, él empieza a acercar su cabeza a la mía cuando llega el momento del golpe cabeza-cabeza. Inmediatamente, cuando llega el momento de acercar la cabeza, retardo el movimiento y es H. quien lo inicia. Se repite la secuencia con H. iniciando el movimiento de acercamiento unas cuantas veces más (p. 59).

Observación 2: A los 0; 9 (25), H. me mira mientras hago ritmos y movimientos con las manos. Solo me mira, luego vuelve a sus objetos. Estiro un pie y lo toco, H. sonríe, lo agarro con las piernas y lo traigo hacia mí y hacia adelante, con las piernas varias veces, mientras canto. H. toma mis pies, los mira; está muy interesado, yo los muevo rítmicamente. H. no suelta los pies, acompaña el movimiento con sus manos. Cuando dejo de moverlos, los mueve él. No nos miramos, estamos de espaldas. Cruzo las piernas. H. se queda mirando mi pie, lo toma y lo mueve hacia los costados. Reinicio el juego. Así varias veces. Toda la escena dura unos cuatro minutos (p. 59).

La observación 1 da cuenta de la anticipación por parte del bebé de la conducta de la adulta mientras ella despliega una *performance* multimodal, interacción en la cual se produce un intenso contacto ocular. Por su parte, la observación 2 muestra el mismo tipo de habilidad, pero en este caso ante una *performance* en la cual (en un momento particular de esta) la única fuente de estimulación que recibe el bebé es la información cinestésica, cuando su cuerpo es movido por los pies de la adulta en ausencia de

contacto ocular. Estos son solo dos ejemplos de cómo se enlaza el modo de participación del bebé durante la interacción y su capacidad de percepción intersensorial.

8.3. El ritmo y la interacción a los 10 meses

Comenzamos este trabajo mencionando que la mayoría de los estudios sobre el desarrollo intersubjetivo temprano se focalizan en la revolución del noveno mes o en el tránsito desde la intersubjetividad primaria a la secundaria. Reclamamos que este no es el único cambio que debe ser explicado. Consecuentemente, hemos dado cuenta de cómo otras habilidades de contacto intersubjetivo pueden vincularse al desarrollo de la percepción intersensorial. Sin embargo, nos parece lícito dar cuenta, en parte, también del tránsito de un modo de intersubjetividad a otro. Una de las habilidades que emerge durante el período de intersubjetividad secundaria es la alternancia visual entre el bebé y el adulto mientras comparten su experiencia con relación a un objeto o evento. Nuevamente, para mostrar el enlace entre el desarrollo de esta habilidad y el desarrollo de la percepción intersensorial, recurriremos a un fragmento de una observación correspondiente a un microanálisis de la interacción entre una madre y su hija de 10 meses en situación de juego social. La observación está tomada de Fogel y DeKoeyer-Laros (2007):

Susan (la hija) mira la mesa y luego mira la cara de Sheryl (la mamá). Susan, levantando las cejas, comienza a golpear. Golpea la mesa con las palmas de sus manos, ella sigue los movimientos de Sheryl con la mirada. Parece que ella quiere iniciar el juego de golpes. Pero Sheryl dice: "Quiero presumir y mostrar cómo puedes hacer pat-a-cake⁷³". Susan se detiene y mira la cara de Sheryl.

⁷³ Canción tradicional inglesa que los adultos suelen cantar a los bebés durante los juegos sociales tempranos, generalmente es acompañada con juego de palmas.

Sheryl comienza a cantar y aplaude “pat-a-cake”. Susan canta y aplaude mirando fijamente a su madre, se detiene y comienza a golpear la mesa de nuevo. Mientras golpea la mesa, alterna su mirada entre sus manos y Sheryl, quien dice: “No, muéstrales cómo aplaudes, ¿de acuerdo?” (p. 80, la traducción es nuestra).

El fragmento de la observación muestra cómo la niña inicia la secuencia de interacción y realiza la alternancia de la mirada entre sus golpeteos y la cara y los golpeteos de la mamá. La interacción descrita no es un juego espontáneo entre la madre y la pequeña hija. Responde al juego que, históricamente, fueron elaborando mientras participaban de la investigación que describió sus conductas. Un ejemplo de las denominadas “*performances* históricas” (Cartero, 2016). Fogel y DeKooyer-Laros (2007) sostienen que la madre andamió la coactuación intersensorial de los golpeteos con el fin de estabilizar los movimientos rítmicos de las manos de ambas participantes permitiendo que la niña alternara su mirada entre sus golpeteos y la conducta y rostro de su mamá. Solo es posible que la niña pueda interactuar con su madre (mientras realiza una coactuación intersensorial) si tiene la capacidad para reconocer que el ritmo de sus golpes sobre la mesa de la sillita alta es el mismo que el de los golpes que realiza su madre. Efectivamente, esto es así debido a que, a los 10 meses, los bebés reconocen relaciones intersensoriales basadas en el ritmo (Martínez, Español e Igoa, en prensa). Nuevamente, pudimos identificar en una situación interactiva el enlace entre el desarrollo de la percepción intersensorial y el desarrollo intersubjetivo temprano.

9. Algunas ideas a futuro

En este trabajo, pusimos en relación los datos empíricos disponibles sobre el desarrollo de la percepción intersensorial (detección y discriminación de patrones rítmicos unimodales y bimodales y reconocimiento de relaciones intersensoriales basadas en el ritmo) con los datos observacionales disponibles sobre la participación del bebé en interacciones en las cuales el adulto le realiza una actuación o una *performance* multimodal modelada rítmicamente. El objetivo fue establecer -en términos dialécticos (Español, 2010; Rivière, 2003)- un modo posible de enlazar el desarrollo de la percepción con el desarrollo intersubjetivo. En tal sentido, creemos que resulta factible comprender cómo el modo de involucramiento del bebé con la actuación o *performance* adulta permitiría el desarrollo de su percepción intersensorial al tiempo que el desarrollo de la percepción intersensorial impactaría en el desarrollo de su involucramiento social. De esta manera, es posible explicar algunos cambios en la conducta interactiva del bebé más allá del tránsito de la intersubjetividad primaria a la secundaria. En estas páginas nos restringimos exclusivamente al ritmo. Sin embargo, si deseamos lograr una comprensión más completa sobre el desarrollo intersubjetivo temprano, deberemos continuar gestando enlaces de este tipo entre el desarrollo de la percepción intersensorial de la sincronía, la duración y la *rate* y la actuación adulta modelada temporalmente por estos mismos factores.

Agradecimientos

Agradezco a Silvia Español, Fernando Rodríguez e Irene Audisio la atenta lectura, los comentarios precisos y los cambios sugeridos que dieron orden y claridad al texto más allá de mis limitaciones.

Referencias bibliográficas

- Allen, T. W., Walker, K., Symonds, L., y Marcell, M. (1977). "Intrasensory and intersensory perception of temporal sequences during infancy". *Developmental Psychology*, 13(3), p. 225.
- Bahrick, L. E. (1988). "Intermodal learning in infancy: Learning on the basis of two kinds of invariant relations in audible and visible events". *Child Development*, 59, pp. 197-209.
- Bahrick, L. E. (2004). "The development of perception in a multimodal environment. En G. Bremner, y A. Slater, (Eds.), *Theories of Infant Development* (pp. 90-120). Malden, MA: Blackwell Publishing.
- Bahrick, L. E. (2010). "Intermodal perception and selective attention to intersensory redundancy: Implications for typical social development and autism". En G. Bremner y T. D. Wachs (Eds.). *The Wiley-Blackwell Handbook of Infant Development: 2nd ed.* (pp. 120-166). Oxford: Blackwell Publishing.
- Bahrick, L. E. y Lickliter, R. (2002). "Intersensory redundancy guides early perceptual and cognitive development". En R. Kail (Ed.), *Advances in Child Development and Behavior*, 30 (pp. 153-187). Nueva York: Academic Press.

- Bahrick, L. E. y Lickliter, R. (2004). "Infants' perception of rhythm and tempo in unimodal and multimodal stimulation: A developmental test of the intersensory redundancy hypothesis". *Cognitive, Affective and Behavioral Neuroscience*, 4, pp. 137-147.
- Bahrick L. E. y Lickliter, R. (2012). "The role of intersensory redundancy in early perceptual, cognitive, and social development". En A. J. Bremner, D. J. Lewkowicz y C. Spence (Eds.), *Multisensory Development* (pp. 183-206). Oxford: Oxford University Press.
- Balmaceda, T. (2018). "Corporalidad y atribución psicológica". En T. Balmaceda y K. Pedace (Comps.), *Temas de filosofía de la mente. Atribución Psicológica* (pp. 15-39). Buenos Aires: SADAFA.
- Baron Cohen, S. (1991). "Precursors to a theory of mind: understanding attention in others". En A. Whiten (Ed.), *Natural Theories of Mind: Evolution, Development and Simulation of Everyday Mindreading* (pp. 233-251). Oxford: Blackwell.
- Baron Cohen, S. (2005). "The Empathizing System: a revision of the 1994 model of the Mindreading System". En B. Ellis y D. Bjorklund (eds.), *Origins of the Social Mind. Evolutionary Psychology and Child Development* (pp. 468-492). Nueva York: Guilford Press.
- Baron Cohen, S., Tager-Flusberg, H., y Lombardo, M. (2013). *Understanding Others Minds. Perspectives from Developmental Social Neuroscience*. Oxford: Oxford University Press.
- Bordoni, M., Español, S. y De Grande, P. (2016). "La incidencia del entonamiento afectivo y la imitación en el involucramiento visual-social temprano". *Avances en Psicología Latinoamericana*, 34(3), pp. 487-503.
- Bremner, A., Lewkowicz, D. y Spence, C. (2012). *Multisensory Development*. Oxford: Oxford University Press.

- Carretero, S. y Español, S. A. (2016). "Multimodal study of adult-infant interaction: a review of its origins and its current status". *Paidéia (Ribeirão Preto)*, 26(65), pp. 377-385.
- Carruthers, P. y Smith, P. K. (Eds.) (1996). *Theories of Theories of Mind*. Cambridge: Cambridge University Press.
- Case, R. (1989). *El desarrollo intelectual. Del nacimiento a la edad madura*. Barcelona: Paidós Ibérica.
- De Jaegher, H. y Di Paolo, E. (2007). "Participatory sense-making". *Phenomenology and the Cognitive Sciences*, 6(4), pp. 485-507.
- Di Paolo, E. A., Rohde, M. y Iizuka, H., (2008). "Sensitivity to social contingency or stability of interaction? Modelling the dynamics of perceptual crossing", *New Ideas in Psychology*, 26(2): pp. 278-294.
- Di Paolo E. A., Rohde M. y De Jaegher H., (2010). "Horizons for the enactive mind: values, social interaction, and play". En J. Stewar, O. Gapenney E. Di Paolo (eds.), *Enaction: Towards a New Paradigm for Cognitive Science*. MIT Press, Cambridge, MA.
- Ekman, P. (1982). *Emotions in the Human Face*. Cambridge: Cambridge University Press.
- Español, S. (2004). *Cómo hacer cosas sin palabras. Gesto y ficción en la infancia temprana*. Madrid: Antonio Machado.
- Español, S. (2008). "Metarrepresentacion e intersubjetividad". En A. Gianella, M. C. González y N. Stigol (Comp.), *Pensamiento, representaciones, conciencia. Nuevas reflexiones* (pp. 113-148). Buenos Aires: Alianza.
- Español, S. (2010). "Interazione precoce. Una prospettiva vygotskijana a partire dagli schemi di Piaget, 25 anni dopo", *Metis*, 17(1) pp. 67-91.

- Español, S. (2014). "La forma repetición-variación: una estrategia para la reciprocidad". En S. Español (Comp.), *Psicología de la música y del desarrollo. Una exploración interdisciplinaria sobre la musicalidad humana* (pp. 157-192). Buenos Aires: Paidós.
- Español, S. y Shifres, F. (2015). "The artistic infant directed performance: A microanalysis of the adult's movements and sounds". *Integrative Psychological and Behavioral Science*, 49(3), pp. 371-397.
- Fantz, R. L. (1961). "The origin of form perception". *Scientific American*, 204(5), pp. 66-73.
- Feldman, R., Magori-Cohen, R., Galili, G., Singer, M. y Louzoun, Y. (2011). "Mother and infant coordinate heart rhythms through episodes of interaction synchrony". *Infant Behavior and Development*, 34(4), pp. 569-577.
- Field, T. (1978). "The three Rs of infant-adult interactions: Rhythms, repertoires, and responsitivity". *Journal of Pediatric Psychology*, 3(3), pp. 131-136.
- Fogel, A., y DeKooyer-Laros, I. (2007). "The developmental transition to secondary intersubjectivity in the second half year: A microgenetic case study". *Journal of Developmental Processes*, 2, pp. 63-90.
- Gallagher, S. (2001). "The practice of mind. Theory, simulation or primary interaction?", *Journal of Consciousness Studies*, 8(5-6), pp. 83-108.
- Gallagher, S. (2008). "Direct perception in the intersubjective context". *Consciousness and Cognition*, 17(2), pp. 535-543.
- Gallagher, S. (2015). "Reuse and body-formatted representations in simulation theory". *Cognitive Systems Research*, 34, pp. 35-43.

- Gallager, S. (2017). "Intercorporeidad y reversibilidad: Merleau-Ponty, emoción, percepción e interacción". En D. Pérez y D. Lawler (Comp.) *La segunda persona y las emociones* (pp. 23-44). Buenos Aires: SADAF.
- Gibson, E. (1969). *Principles of Perceptual Learning and Development*. Nueva York: Appleton-Century-Crofts.
- Gogate, L. J., Bolzani, L. H. y Betancourt, E. A. (2006). "Attention to maternal multimodal naming by 6- to 8-month-old infants and learning of word-object relations". *Infancy*, 9(3), pp. 259-288.
- Gómez, J. C. (1996). "Second person intentional relations and the evolution of social understanding". *Behavioral and Brain Science*, 19(1), p. 129.
- Gómez, J.C. (2005). "Joint Attention and the Notion of Subject: Insights from Apes, Normal Children, and Children with Autism". En N. Eilan, C. Hoerl, T. McCormack y J. Roeleser (Eds.), *Joint Attention: Communication and Other Minds. Issues in Philosophy and Psychology* (pp. 65-84). Oxford: Oxford University Press.
- Gómez, J. C. (2008). "The evolution of pretense: From intencional availability to intencional non-existence". *Mind and Language*, 23(5), pp. 586-606.
- Gómez, J.C. (2009). "Embodying meaning: Insights from primates, autism, and Brentano". *Neural Networks*, 22, pp. 190-196.
- Gómez, J.C. (2010). "L'evoluzione del mentalismo intenzionale: della disponibilitá intenzionale alle intenzioni sul nulla". *Metis*, 17: pp. 7-27.
- Gómez, J. C. y Núñez, M. (1998). "La mente social y la mente física: desarrollo y dominios de conocimiento". *Infancia y aprendizaje*, 21(84), pp. 5-32.

- Gomila, A. (2003). "La perspectiva de segunda persona de la atribución mental". En A. Duartey E. Rabossi (comp.), *Psicología cognitiva y filosofía de la mente* (pp. 195-218). Buenos Aires: Alianza Editorial.
- Gomila, A. y Pérez, D. (2017). "Lo que la segunda persona no es". En D. Pérez y D. Lawler (Comp.), *La segunda persona y las emociones* (pp. 275-297). Buenos Aires: SADAFA.
- Harris, P.L. (1991). "The work of imagination". En A. Whiten (Ed.), *Natural Theories of Mind* (pp. 283-304). Oxford: Basil Blackwell.
- Harris, P.L. (1992). *Los niños y las emociones*. Madrid: Alianza.
- Hobson, R. P. (1995). *El autismo y el desarrollo de la mente*. Madrid: Alianza.
- Hubley, P. y Trevarthen C. (1979). "Sharing a task in infancy". En I. Uzgiris (Ed.), *Social Interaction during Infancy: New Directions for Child Development*. Vol. 4 (pp. 57-80). San Francisco: Jossey Bass.
- Jaffe, J., Beebe, B., Feldstein, S., Crown, C. L., Jasnow, M. D., Rochat, P. y Stern, D. N. (2001). "Rhythms of dialogue in infancy: Coordinated timing in development". *Monographs of the Society for Research in Child Development*, 66(2), p. 149.
- Kaye, K. (1986). *La vida mental y social del bebé. De cómo los padres crean personas*. Barcelona: Paidós.
- Koester, L. S., Papoušek, H. y Papoušek, M. (1989). "Patterns of rhythmic stimulation by mothers with three-month-olds: A cross-modal comparison". *International Journal of Behavioral Development*, 12(2), pp. 143-154.

- Leslie, A. (1988). "Some implications of pretense for mechanisms underlying the child's theory of mind". En J. W. Astington, P. L. Harris y D. P. Olson (Eds.), *Developing Theories of Mind* (pp. 19-46). Cambridge: Cambridge University Press.
- Lewkowicz, D. J. (1994). "Development of intersensory perception in human infants". En D. J. Lewkowicz y R. Likliter (Eds.), *The Development of Intersensory Perception. Comparative Perspectives* (pp. 165-203). Nueva Jersey: Lawrence Erlbaum Associates.
- Lewkowicz, D. J. (2000). "The development of intersensory temporal perception: An epigenetic systems/limitations view". *Psychological Bulletin*, 126(2), pp. 281-308.
- Lewkowicz, D. J. (2003). "Learning and discrimination of audiovisual events in human infants: the hierarchical relation between intersensory temporal synchrony and rhythmic pattern cues". *Developmental Psychology*, 39(5), p. 795.
- Lewkowicz, D. J. y Marcovitch, S. (2006). "Perception of audiovisual rhythm and its invariance in 4- to 10-month-old infant". *Developmental Psychobiology*, 48(4), pp. 288-300.
- Martínez, M. (2010). *De la intersubjetividad primaria a la secundaria: ¿qué percibe el bebé cuando mira al adulto?* Tesis de Maestría sin publicar. FLACSO-UAM.
- Martínez, M. (2011). "Intersubjetividad y teoría de la mente. Un mapa para comprender sus relaciones y las diversas formas que adopta su investigación". *Psicología del Desarrollo*, 1(2), pp. 9-29.
- Martínez, M. (2016). "Desarrollo de la percepción inter-sensorial. Aspectos, metodológicos, empíricos y conceptuales". En E. Huairé Inacio, Á. Elgier y G. Clerici

- (Comps.), *Pensar la niñez: psicología del desarrollo desde una perspectiva americana* (pp. 39-59). Lima: Editora y Librería Jurídica Grijley EIRL.
- Martínez, M., Español. S. e Igoa, J. M. (2018). "Reconocimiento de relaciones intersensoriales basadas en el ritmo a los 4, 7 y 10 meses". *Epistemus*.
- Martínez, M., Español, S. e Igoa, J. M. (en preparación). "Reconocimiento de relaciones intersensoriales basadas en la duración".
- Martínez, I. C. (2007). "La composicionalidad de la performance adulta en la parentalidad intuitiva". En M. de la P. Jacquier y A. Pereira Ghiena (Eds.) *Música y Bienestar Humano. Actas de la VI Reunión de SACCoM* (pp. 25-34). Buenos Aires: SACCoM.
- Martínez, I. C. (2008). "La composición temporal del habla, el canto y el movimiento en la musicalidad de las interacciones tempranas adulto-infante". En M. de la P. Jacquier y A. Pereira Ghiena (eds.), *Objetividad-Subjetividad y Música. Actas de la VI Reunión de SACCoM* (pp. 73-82). Rosario: SACCoM.
- Martínez, I. C. (2014). "La base corporeizada del significado musical". En S. Español (comp.), *Psicología de la música y del desarrollo. Una exploración interdisciplinaria sobre la musicalidad humana* (pp. 71-110). Buenos Aires: Paidós.
- Martínez, I.C., Español, S y Pérez, D. (2018). "The interactive origin and the aesthetic modelling of image-schemas and primary metaphors". *Integrative Psychological and Behavioral Science*, 52(4), pp. 646-671.
- McNeil, W. H. (1995). *Keeping Together in Time: Dance and Drill in Human History*. Cambridge: Harvard University Press.

- Melinder, A., Forbes, D., Tronick, E., Fikke, L. y Gredebäck, G. (2010). "The development of the still-face effect: Mothers do matter". *Infant Behavior and Development*, 33(4), pp. 472-481.
- Miall, D. S., y Dissanayake, E. (2003). "The poetics of baby-talk". *Human nature*, 14(4), pp. 337-364.
- Moreno-Núñez, A., Rodríguez, C. y Del Olmo, M. J. (2015). "The rhythmic, sonorous and melodic components of adult-child-object interactions between 2 and 6 months old". *Integrative Psychological and Behavioral Science*, 49(4), pp. 737-756.
- Ospina, V. y Español, S. (2014). "El movimiento y el sí mismo. En S. Español (Ed.), *Psicología de la música y del desarrollo. Una exploración interdisciplinaria sobre la musicalidad humana* (pp. 111-155). Buenos Aires: Paidós
- Papoušek, H. (1996). "Musicality in infancy research: biological and cultural origins of early musicality". En I. Deliège y J. Sloboda (Eds), *Musical Beginnings. Origins and Development of Musical Competence* (pp. 37-55). Oxford: Oxford University Press.
- Papoušek, M. (1996). "Intuitive parenting: A hidden source of musical stimulation in infancy". En I. Deliège y J. Sloboda (Eds.), *Musical Beginnings. Origins and Development of Musical Competence* (pp. 88-112). Oxford: Oxford University Press.
- Papoušek, M. y Papoušek, H. (1981). "Musical elements in the infant's vocalizations: Their significance for communication, cognition and creativity". En L. P. Lipsitt (Ed.), *Advances in Infancy Research*, Vol 1. (pp. 63-224). Nueva Jersey: Ablex Norwood.
- Pérez, D. (2013). *Sentir, desear, creer*. Buenos Aires: Prometeo.

- Pérez, D. y Gomila, A. (2018). "La atribución mental y la segunda persona". En T. Balmaceda y K. Pedace (comps.), *Temas de filosofía de la mente. Atribución psicológica* (pp. 69-98). Buenos Aires: SDAF.
- Perinat, A. (1993). *Comunicación animal, comunicación humana*. Madrid: Siglo XXI.
- Perner, J. (1994). *Comprender la mente representacional*. Barcelona: Paidós.
- Phillips-Silver, J., Aktipis, C. A. y Bryant, G. A. (2010). "The ecology of entrainment: Foundations of coordinated rhythmic movement". *Music Perception: An Interdisciplinary Journal*, 28(1), pp. 3-14.
- Pickens, J. N. y Bahrick, L. E. (1995). "Infants' discrimination of events on the basis of rhythm and tempo". *British Journal of Developmental Psychology*, 13, pp. 223-236.
- Provasi, J., Anderson, D. I., y Barbu-Roth, M. (2014). "Rhythm perception, production, and synchronization during the perinatal period". *Frontiers in Psychology*, 5, p. 1048.
- Reddy, V. (1996). "Omitting the second-person in social understanding". *Behavioral and Brain Sciences*, 19(1), pp. 140-141.
- Reddy, V. (2008). *How Infants Know Minds*. Harvard: Harvard University Press.
- Reddy, V. y Morris, P. (2009). "Participants Don't Need Theories: Knowing Minds in Engagement". En I. Leudar y A. Costall (Eds.), *Against Theory of Mind* (pp. 91-107). Londres: Palgrave Macmillan.
- Rivière, A. (2003). "Interacción precoz. Una perspectiva vygotskiana a partir de los esquemas de Piaget". En M. Belinchón, A. Rosa, M. Sotillo y I. Marichalar (comp.), *Ángel Rivière. Obras escogidas*, Vol. II (pp. 109-142). Madrid: Panamericana.

- Rochat, P. (2004). *El mundo del bebé*. Madrid: Morata.
- Roszmanith, N., Costall, A., Reichelt, A. F., López, B., y Reddy, V. (2014). "Jointly structuring triadic spaces of meaning and action: Book sharing from 3 months on". *Frontiers in Psychology*, 5, p. 1390.
- Sarriá, E. y Rivière, Á. (1991). "Desarrollo cognitivo y comunicación intencional preverbal: un estudio longitudinal multivariado". *Estudios de Psicología*, 46, pp. 35-52.
- Scotto, C. (2002). "Interacción y atribución mental: la perspectiva de la segunda persona". *Análisis filosófico*, 22(2), pp. 135-155.
- Scotto, C. (2017). "Lo que el aprendizaje del lenguaje revela sobre el lenguaje (y sobre la cognición social)". En D. Pérez y D. Lawler (Comps.), *La segunda persona y las emociones* (pp. 87-140). Buenos Aires: SADAF.
- Schaffer, R. H. (1977). "Early interactive development". En H. R. Schaffer (Ed.), *Studies in Mother-Infant Interaction: Proceedings of the Loch Lomond Symposium* (pp. 3-16). Londres: Academic Press.
- Schaffer, R. H. (1983). *El desarrollo de la sociabilidad*. Madrid: Aprendizaje Visor.
- Schögler, B. y Trevarthen, C. (2007). "To sing and dance together". En S. Bråten (Ed.), *On Being Moved: From Mirror Neurons to Empathy* (pp. 281-302). Ámsterdam/Philadelphia: John Benjamins.
- Shifres, F. (2014). "Algo más sobre el enlace entre la infancia temprana y la música: el poder expresivo del rubato". En S. Español (comp.), *Psicología de la música y del desarrollo. Una exploración interdisciplinaria sobre la musicalidad humana* (pp. 21-70). Buenos Aires: Paidós.
- Stern, D. N. (1983). *La primera relación madre-hijo*. Madrid: Morata.

- Stern, D. N. (1991). *El mundo interpersonal del infante. Una perspectiva desde el psicoanálisis y la psicología evolutiva*. Buenos Aires: Paidós.
- Stern, D. N., Beebe, B., Jaffe, J. y Bennett, S. L. (1977). "The infant's stimulus world during social interaction". En H. R. Schaffer (Ed.), *Studies in Mother-Infant Interaction: Proceedings of the Loch Lomond Symposium* (pp. 177-202). Londres: Academic Press.
- Stern, D., Hofer, L., Haft, W. y Dore, J. (1985). "Affect attunement: The sharing of feeling states between mother and infant by means of intermodal fluency". En T. Field y N. Fox (Eds.), *Social Perception in Infants* (pp. 249-268). Norwood, NJ: Ablex.
- Stevanovic, M. y Koski, S. E. (2018). "Intersubjectivity and the Domains of Social Interaction: Proposal of a Cross-Sectional Approach". *Psychology of Language and Communication*, 22(1), pp. 39-70.
- Tomasello, M. (2013). *Los orígenes de la comunicación humana*. Barcelona: Katz Editores.
- Trevarthen, C. (1984). "Emotion in infancy: Regulators of contacts and relationships with persons". En K. Scherer y P. Ekman (Eds.), *Approaches to Emotions* (pp. 129-157). Hillsdale: Erlbaum.
- Varela, F., Thompson, E. y Rosch, E. (1994). *De cuerpo presente. Las ciencias cognitivas y la experiencia humana*. Barcelona: Gedisa.
- Verhage, (2008). *The Rhythm of Embodied Encounters: Intersubjectivity in Merleau-Ponty's Phenomenology*. Tesis de doctorado sin publicar. McGill University.
- Walker-Andrews, A. S. (1997). "Infants' perception of expressive behaviors: differentiation of multimodal information". *Psychological Bulletin*, 121(3), p. 437.

Wellman, H. (2017). *La construcción de la mente. Cómo se desarrolló la teoría de la mente*. Santiago de Chile: Ediciones UC.

Competencia social y reconocimiento de emociones a partir de expresiones faciales y lenguaje corporal en sujetos del espectro autista

Uso de plataformas virtuales

XILENIA MARÍA CARRERAS (SECYT – UNC)

1. Introducción

La comprensión de las emociones mediada por el tipo de interacciones socioemocionales comienza en la diá-dia madre-hijo de la infancia temprana, en las proto-conversaciones, basadas en la comunicación corporal y la acción conjunta que suceden entre ellos. Esta compren-sión del mundo social supone tener en cuenta la conducta de los demás situándola en su contexto y ser capaz de com-prender sus estados emocionales (Segovia Cuellar, 2017). Esta capacidad, a la que Trevarthen (1998) ha denominado “intersubjetividad primaria” es entendida como la capaci-dad de entender las acciones de los demás en contexto, ya que percibimos a los demás como agentes intencionales y tenemos acceso a sus emociones a través de la natura-leza comunicativa de su cuerpo (Segovia Cuellar, 2017). El enamoramiento o “enganche”, visible en los primeros meses de edad, es lo que permite esa atención privilegiada

de los niños pequeños hacia sus cuidadores. Este “enganche” es crucial en los primeros años de vida ya que posibilita a los niños sentir confianza, intimidad y calidez, pilares para el desarrollo socioemocional. En la medida en que las interacciones resultan placenteras, habrá motivación para atender al entorno (escuchar y mirar a un cuidador, atender a olores y otras sensaciones del entorno). También les permitirá regularse a sí mismos en base a la confianza y calidez con un cuidador, tener una intención comunicativa, figurarse cierta orientación espacial (dónde está mamá o papá, por ejemplo) y mantener la motivación en tareas de imitación y otros procesos de aprendizaje más sofisticados (Greenspan y Wieder, 2006).

Las personas con autismo -una condición neurobiológica de variedad muy heterogénea que afecta el desarrollo temprano- presentan gran variedad de síntomas, como dificultades en la comunicación, en la conducta intencional, en el reconocimiento y expresión de las emociones a partir de expresiones faciales, en el habla y el lenguaje corporal (Enticott *et al.*, 2014). Si existen alteraciones de este tipo (observables en las primeras etapas del desarrollo), cabría esperar alteraciones en el desempeño social, en general, y en la lectura de emociones más sofisticadas.

Atendiendo al tipo de desafíos presentes en esta población, se han creado soportes y ayudas para las personas con necesidades educativas especiales mediante el uso de las nuevas tecnologías. Dado que la expresión de emociones, estados de humor y de actitudes es parte central de la comunicación en contextos sociales, a partir de modelos de interacción humana en el campo de la robótica se han diseñado sistemas como el Rachel, un tipo de agente conversacional corporizado (*Embodied Conversational Agent*, ECA según sus siglas en inglés) (Kirby *et al.*, 2010).

Teniendo en cuenta que el uso de las nuevas tecnologías es cada vez más frecuente y aceptado en el ámbito educativo y clínico, nos proponemos aquí delinear algunos requisitos que tendrían que cumplir para servir de apoyo a la experiencia teniendo en cuenta un enfoque corporeizado.

2. El enfoque corporeizado y la cognición social corporizada

En el cognitivismo clásico, se parte de la metáfora computacional para explicar los procesos del pensamiento y la acción, que se entienden como un conjunto de operaciones computacionales formales ejecutadas desde el sistema cognitivo. Desde esta perspectiva, el pensamiento está separado de los procesos sensoriomotores.

Para el enfoque corporeizado, que se presenta como una alternativa a esta perspectiva, el sistema sensoriomotor es constitutivo de la cognición, lo que le da un lugar central al cuerpo y la percepción (Sanabria, 2017). Parte de nuestro sistema conceptual y algunos aspectos de nuestro lenguaje están estructurados por las características de nuestro cuerpo. Desde esta concepción, la propia geografía corporal determina el tipo de experiencias posibles, y esa es la forma que adquiere nuestra cognición. No es lo mismo la experiencia de bipedia, con dos ojos, uno a la par del otro con manos de pulgar contrapuesto, por ejemplo, que si nuestra geografía/constitución corporal nos diera otra experiencia posible (Pérez y Lawler, 2017).

Desde este punto de vista, lo que hacemos cuando interactuamos con los otros es interpretar sus acciones en determinadas situaciones contextualizadas (cognición social corporizada). No haría falta tener una representación del estado mental del otro, ya que accedemos a él

directamente a través de su comportamiento, su lenguaje corporal y su gestualidad. Lo que hacemos es “leer cuerpos” antes que leer mentes. La percepción social directa implica el acceso directo a los estados emocionales de los otros, es decir:

Estamos informados del gozo de los otros en su risa, de su dolor y sufrimiento en sus ojos, de su pena en el rubor, de su súplica en las manos extendidas, de su amor en su mirada de afecto, de su rabia en el crujir de sus dientes, y el tono de sus pensamientos en el volumen y sonido de sus palabras (Scheler, 1954, p. 260, en Segovia Cuellar, 2017, p. 19).

Esa percepción directa conforma “una corporalidad mutua’ desde ahí, captamos las intenciones de los otros, sus ideas y creencias. Mediante la observación directa, tenemos acceso a otras mentes, y por eso es por lo que se habla del entendimiento social práctico que se da a partir de las interacciones significativas con otros. En este sentido:

Una teoría que se centra en la interacción corporal también ayuda a explicar la aparición de nuevas formas de entendimiento y práctica social como la intencionalidad compartida. La intencionalidad compartida supone la aparición de una motivación y una comprensión social extensiva, al parecer únicamente humana, que permite conductas cooperativas y colaborativas complejas como la atención conjunta, la comunicación declarativa, o la colaboración conjunta (Segovia Cuellar, 2017, p. 365).

Desde este abordaje corporizado, en los siguientes apartados revisaremos la noción de sensibilidad a los estímulos sociales y la contingencia social (atender a rostros humanos en una situación de interacción social), la atención conjunta y la aparición de los primeros gestos, como los de mostrar y pedir, en infantes de desarrollo típico, habilidades necesarias para la competencia social.

3. Cognición social, lenguaje corporal y gestualidad en el desarrollo

La sensibilidad a la contingencia social, observable desde los primeros meses de vida, ha sido estudiada por experimentos clásicos en los que se expone a niños muy pequeños a respuestas recíprocas versus neutras (cara estática). En el clásico experimento presentado inicialmente por Tronick, Adamson, Als y Brazelton (1978), en el que se observan interacciones uno a uno entre infantes de 20 semanas de edad y sus madres, se les pide a las madres que interrumpan la interacción, volteando la cara, y que reaparezcan con una expresión neutra, haciendo caso omiso a los intentos comunicativos del niño durante tres minutos. Este paradigma experimental del *stillface* mostró que los niños de desarrollo típico suelen variar el estado afectivo que tenían, de mutua reciprocidad, por uno neutro o de malestar ante las caras neutras, en especial tras varios intentos fallidos de convocar la atención del adulto⁷⁴.

El andamiaje dado con el otro, con el afecto como vía privilegiada, favorece la atención conjunta, la cual puede ser considerada como una habilidad necesaria para una adecuada interacción social más compleja. Se entiende por “atención conjunta” a “la capacidad de coordinar la atención con otra persona en relación con un objeto o acontecimiento” (Montagut Asunción, D’Ocon Giménez y Pastor Cerezuela, 2017, p. 375). Esta tiene una función social e instrumental (por ejemplo, conseguir un objeto) y aparece alrededor de los doce meses.

La gestualidad como tal aparece entre los 9 y 13 meses en forma de conductas como dar, mostrar y pedir. Posteriormente, alrededor de los 24 meses, se incorporan gestos

⁷⁴ También aparece en el desarrollo lo que se denomina como ‘referenciación social’ donde los niños suelen atender a la reacción del adulto ante un estímulo novedoso que puede ser interpretado como peligroso.

más complejos, como el de señalar a la distancia. Las primeras interacciones sociales ocurren de modo muy temprano en el desarrollo, también favorecidas por la tendencia de los bebés pequeños a mantener los ojos abiertos y atender, en los casos de desarrollo típico, a estímulos sociales (una cara, un tono de voz) antes que a estímulos no sociales.

En el autismo se presenta una dificultad para codificar algunas de estas señales implícitas, por ejemplo, un déficit inicial para atender a estímulos sociales⁷⁵ y para el reconocimiento de las emociones y los estados mentales en los demás. Estas fallas en la atención conjunta se observan tanto al iniciar como al responder a intentos comunicativos, por lo que la comunicación se vería afectada por déficits en estas primeras experiencias de reciprocidad.

4. Desafíos presentes en sujetos con CEA

Las personas con trastorno del espectro autista (TEA, para la nomenclatura actual del *DSM-V*, y con condición del espectro autista, CEA., según la jerga actual impulsada por el paradigma de la neurodiversidad⁷⁶) presentan un déficit sociocomunicativo, así como también patrones restrictivos o repetitivos del comportamiento, de los intereses o las actividades. Según el nivel de estas dificultades, se precisa también en el diagnóstico la necesidad de apoyo (necesidad de “ayuda”, “ayuda notable” y “ayuda muy notable”).

⁷⁵ Desde los modelos de la cognición social, se ha señalado la importancia de la atención conjunta como requisito para la comunicación social, rasgo alterado en el caso del autismo (véase por ejemplo Osterling J. A., Dawson G., Munson J. A., 2002).

⁷⁶ Véanse, por ejemplo, los comunicados de divulgación y promoción de la salud destinada a niños y niñas del espectro en la ONG PANAACEA (cfr. <https://www.panaacea.org/>).

Algunos indicadores del diagnóstico temprano son: 1) conductas de atención conjunta menos frecuentes y/o eficaces; 2) comunicación y percepción gestual más limitada; 3) menor rendimiento en tareas lingüísticas; 4) diferente percepción de rostros y de otros estímulos sociales; 5) sonrisa social más débil; 6) menor presencia de juego simbólico.

Todos estos indicadores son observables, incluso, en hermanos de niños con TEA en comparación con los hermanos de niños y niñas con desarrollo típico. También se observan, en el caso del autismo, rasgos clínicos como contacto visual, sonrisa social y sonrisa social responsiva escasos, presencia de manierismos o movimientos inusuales de los ojos, las manos o la boca, tendencia a dar saltitos, aletear o balancearse, uso inusual de los objetos, alineamiento de sus juguetes, rechazo muy fuerte a ciertos olores, sabores o texturas, búsqueda inusual de ciertos estímulos (por ejemplo, visuales), obsesión por ciertos patrones repetitivos, inflexibilidad, llanto o risa "sin motivo aparente", estar "colgados" mirando alguna zona del ambiente sin que haya ningún estímulo aparentemente llamativo, deambular sin motivo, dificultades en ciertas tareas de tipo ejecutivas (como ideación y planificación y ejecución motora), inhibición, retraso significativo o ausencia de lenguaje, parecer "sordos", etc. (véase, por ejemplo, las escalas estandarizadas de cribado como el CHAT, Cars, o la escala australiana de síndrome de Asperger).

En relación con el autismo, se ha señalado que tienen dificultades en el reconocimiento de emociones y estados mentales tanto propios como ajenos, fallas en la atención conjunta, menor presencia de gestos como mostrar, pedir y señalar, y uso mecánico del cuerpo del otro para obtener algo (llevar de la mano). En los experimentos de *Stillface*, que mencionamos en el apartado anterior, los niños con autismo (y sus hermanitos) tenían una tendencia a man-

tener la expresión neutra aun cuando los adultos reiniciaban la interacción. En los niños⁷⁷ con autismo y hermanos, se identifican retrasos en los rasgos típicos de gestualidad desde el primer año. También se ha encontrado que presentan una tendencia a atender más a estímulos no sociales (por ejemplo, cara invertida) que a estímulos sociales (por ejemplo, un rostro humano). Esto se manifiesta, por ejemplo, en que prestan un menor tiempo de atención hacia el cuidador o hacia un estímulo de rostro, y un mayor tiempo de atención hacia los objetos (Bhat, Galloway y Landa, 2010; Droucker, Curtin y Vouloumanos, 2013).

A nivel empírico, se han encontrado alteraciones en el reconocimiento de algunas expresiones faciales, como el miedo o el enfado, así como en el reconocimiento de caras, porque son menos atractivas para ellos que para otros niños dado que presentan dificultades en integrar las expresiones faciales en una entidad coherente. También se ha encontrado que tienen una limitada capacidad para percibir las emociones ajenas (Begeer, Koot, Rieffe, Terwogt y Stegge, 2008, en García Arnaz *et al.*, 2018).

Si el reconocimiento de las emociones y las primeras interacciones relacionales se ven facilitadas por la interacción cara a cara, y si existe una alteración desde momentos tempranos del desarrollo en el caso del autismo (véase, por ejemplo, Perrault *et al.*, 2018), se verá afectado, entonces, el desarrollo posterior de habilidades más complejas, ya que las situaciones que promueven el intercambio social se encuentran con “ruido”, alteradas o disminuidas.

⁷⁷ Estadísticamente es más frecuente encontrar a niños dentro del espectro que a niñas. Según la teoría del cerebro masculino extremo, el autismo sería una exacerbación de la tendencia a la sistematización que se observa mayoritariamente en los hombres (véase por ejemplo Baron-Cohen, Lombardo, Auyeung, Ashwin, Chakrabarti, Knickmeyer, 2011).

Es posible que algunas personas que no se sienten cómodas ante una situación social prefieran evitarla, por lo que un apoyo a la experiencia con plataformas virtuales, como las apps, puede enriquecer y favorecer este tipo de aprendizaje. Este tipo de apoyo podría ser la vía de acceso privilegiada para los intereses comúnmente expresados en la población con autismo de identificar claves sociales como la lectura de gestos, el reconocimiento de expresiones faciales y lenguaje corporal en general, requisitos previos para una buena lectura de situaciones sociales. Teniendo en cuenta las dificultades de trasladar habilidades que han sido enseñadas de manera directa y segmentadas en unidades discretas (saludar, respetar turnos, mirar a los ojos, mantener una distancia óptima, etc.) a ambientes interactivos en tiempo real, postulamos algunos requisitos y lineamientos generales (como, por ejemplo, la construcción desde un modelo de diseño colaborativo) para el diseño y utilización de aplicaciones que puedan enriquecer la experiencia o volverla socialmente significativa.

Hay que desarrollar las ventajas que ofrecen las tecnologías de la comunicación en situaciones donde los procesos de comunicación resultan un reto para la comunidad educativa, en concreto, para atender a las necesidades de comunicación de las personas implicadas en los procesos de enseñanza-aprendizaje, en los casos donde la comunicación puede estar influida por las características de las personas que se están comunicando. La propuesta de un modelo de intervención con estudiantes autistas supone tomar elementos esenciales de procesos de comunicación y motivación que aportan las TIC (Medina Rivilla y Rodríguez Serna, 2016, p. 1).

Teniendo en cuenta esta sugerencia, en el siguiente apartado haremos un repaso por algunas de las propuestas realizadas desde plataformas virtuales para personas con CEA.

5. Apoyo a la experiencia mediante plataformas virtuales

El desarrollo de nuevas tecnologías ha sido empleado para el entrenamiento de habilidades de distinto tipo, tanto en sujetos con desarrollo típico, como en aquellos con necesidades educativas especiales (Calzadilla, 2002; Ramírez y Cassinero, 2014; Del Moral Pérez, Bellver Moreno y Guzmán Duque, 2018). En el caso del autismo, se han desarrollado plataformas para fines clínicos y educativos. Veremos aquí algunos ejemplos de plataformas de distinto tipo (aplicaciones en su mayoría, y dos casos de robótica como un medio de información ampliada) utilizadas en el reconocimiento de emociones y expresiones faciales. Señalaremos algunas de las ventajas y desafíos presentes en dicho campo, para recuperar luego dos experiencias de diseño colaborativo aplicadas a habilidades prosociales que parecen ser superadoras en cuanto a su eficacia e impacto terapéutico.

Hasta aquí nos hemos focalizado en la lectura de expresiones faciales aplicadas a la lectura de emociones porque, en principio, parecen habilidades de lectura corporal, situadas y que precisan de considerar información tanto estática como dinámica. García Arnáz *et al.* (2018), por ejemplo, implementaron una aplicación denominada Emocionatest como un modo de evaluar competencia emocional en niños autistas, siguiendo como criterio de evaluación hitos del desarrollo típico de niños entre 3 y 12 años. La aplicación buscaba diferenciar errores en la atribución de emociones de valencia diferente (alegría versus tristeza) y de valencia similar (tristeza, enfado). Se pretendía evaluar el rendimiento de niños con CEA en comparación con lo esperable para su edad. En dicho estudio participaron 33 niños con autismo de entre 4 y 12 años. Las tareas de reconocimiento emocional se organizaban

según niveles de dificultad, y estos eran: 1) reconocimiento de expresión facial (identificar la expresión esperada para el contexto entre seis expresiones faciales presentadas); 2) construcción de expresiones faciales a partir de la región ocular y zona de la boca; 3) identificación de emociones en situaciones simples (atribución emocional); 4) reconocimiento emocional a partir de historias o situaciones problemáticas (reconocimiento cognoscitivo); y 5) diferenciación entre emociones reales y fingidas.

Los investigadores encontraron que había un mayor porcentaje de aciertos en el primer nivel de dificultad, seguido por el nivel 3 y luego el nivel 4. Finalmente, el nivel de dificultad encontrado era menor en los niveles 2 y 5. En cuanto al contenido, las mayores dificultades se encontraron en el reconocimiento de la expresión facial que correspondía al miedo, ya que tenía tendencia a confundirse con sorpresa, confusión que también se observó en el nivel 2, donde se confundía sorpresa con felicidad. En la construcción de expresiones faciales a partir de elementos aislados, los desafíos estaban en la zona de los ojos, y no tanto en la forma de la boca para determinada expresión emocional. En el nivel 3 –recordemos que los sujetos tenían que identificar un estado emocional a partir de historias simples–, se encontró que había confusión entre las historias de felicidad y las de sorpresa, y que las segundas se confunden también con las de miedo. En el último nivel, el de reconocimiento cognoscitivo de las emociones, encontraron que este dependía de la información contextual y valencia emocional; por ejemplo, resultaba facilitador que primero estuviera involucrada una emoción positiva, como felicidad, y luego una emoción valorada negativamente, como enojo o tristeza. En cuanto a la distinción entre emociones reales y las fingidas, los participantes tuvieron muchas dificultades en cumplir esa tarea, y acertaron solo un 11,6 % de

intentos. Como punto fuerte de esta experiencia, se puede señalar que incluyeron grupos de control y un aumento progresivo de dificultad, lo cual permite estimar el rendimiento real y esperable, así como también el nivel de complejidad donde se presenta la dificultad.

Otro proyecto, denominado AUTHIC (Castillo Aguilar *et al.*, 2016), se propuso favorecer la comprensión e interpretación de expresiones faciales asociadas a estados emocionales en población autista. Al igual que los anteriores, en este proyecto se trabajó con pictogramas o fotografías de diferentes emociones y fases, en el marco de tareas de diferente complejidad. Primero, se pretendía enseñar por asociación cada emoción con su respectiva expresión facial y se les daba oportunidades a los usuarios de asignar una emoción a su expresión por cuenta propia. En la fase 2 se les pedía a los sujetos que identificaran elementos individuales de un gesto. Por su parte, la fase 3 correspondía a una tarea de identificación de emociones a partir de una situación, y, finalmente, la cuarta fase refería a tareas de imitación de una expresión facial. Aquí vemos que predomina el aprendizaje por asociación e identificación de expresiones estáticas, atendiendo al detalle. Los desafíos más grandes de este tipo de propuestas son la generalización de la habilidad a contextos naturales.

Empleando otro tipo de diseños, Crisol Moya y Domingo Martos (2015) hicieron uso del soporte denominado EXPRESATE, destinado para sujetos con síndrome de Asperger⁷⁸. Este soporte estaba basado en el modelo ASSURE, un tipo de diseño progresivo en el que los docentes pueden diseñar y desarrollar un ambiente de enseñanza-aprendizaje más adecuado para los sujetos. Como rasgos

⁷⁸ Se trata de un subtipo clínico dentro del cuadro “trastorno generalizado del desarrollo”, nomenclatura empleada para los sujetos del espectro autista en el DSM-IV.

principales, EXPRESATE permitía contemplar los estilos de aprendizaje y la participación activa de los usuarios, pudiendo seleccionar objetivos, estrategias, tecnologías, medios y materiales, y organizar el escenario de aprendizaje, evaluación y revisión de resultados. La aplicación estaba pensada para abordar dificultades emocionales a partir de la creación de proyectos que incluían distintos tipos de actividades, como rompecabezas o palabras cruzadas. La creación y diseño de proyectos permitía seleccionar el tipo de imágenes y pictogramas que utilizar (por ejemplo, una caricatura, una representación gráfica, como un emoticón, una fotografía, etc.) y, en un segundo momento, el tipo de operación que realizar (por ejemplo, identificar una emoción, reconocer expresiones faciales o atribuir la causa a dicho estado emocional). Cabe señalar, a modo de crítica, que esta propuesta, cuyos fines han sido exclusivamente educativos, no ha sido evaluada en términos de tamaño del efecto, ni empleaba combinación de grupos de pares, y el rol docente era pensado como un rol de apoyo, en que describía y situaba ciertas expresiones faciales a través de la imitación (por ejemplo, “Esta es mi cara de contento, ahora tú”), por lo cual no estaban involucradas situaciones reales, que incluyen información dinámica que requiere ser integrada para una lectura adecuada del contexto.

Por otro lado, Hourcade, Miller, Huebner y Liang (2013) emplearon una aplicación para niños con autismo con el fin de fomentar interacciones sociales mediante actividades expresivas, creativas y colaborativas, y compararon su comportamiento en actividades sociales con y sin el soporte de la aplicación. El resultado fue que encontraron que el uso de esta *app* estaba asociado al incremento de la comunicación verbal, la interacción física y acotaciones de soporte a otros niños. También se tuvo en cuenta la presencia de comentarios de displacer, comportamiento

atípico socialmente inapropiado (por ejemplo, invasión del espacio personal, tono de voz inapropiado, interrupciones, etc.) y el tiempo de ejecución de cada tarea. En el primer bloque de tareas, los sujetos debían incluir dentro de una historia breve dibujos de rostros, y seleccionar posibles escenas siguientes con el objetivo de fomentar la creatividad, las habilidades de motricidad fina, la actitud de compartir (mostrar) y de colaboración. Otro bloque de tareas incluía una actividad de producción colaborativa de recursos sonoros/musicales. La tercera actividad también era tarea colaborativa de seguimiento de un objetivo, que consistía en la realización de un recorrido de líneas rectas mediante la unión de círculos dispuestos de manera multiforme en la pantalla; aquí también se buscaba favorecer la motricidad fina y fomentar la actitud de compartir. Finalmente, la última actividad era de modelado de expresiones faciales, empleando fotografías de adultos que ellos conocían (modelando y deformando, por ejemplo, lo que entretenía mucho a los participantes). Aquí los sujetos fueron ocho niños de entre 10 y 14 años. Este proyecto de habilidades combinadas (motricidad fina y promoción de actitudes prosociales) presentaba la novedad de haber incluido las fotos de los adultos que los niños conocían. Por otro lado, también se promovía la interacción cara a cara, mientras interactuaban con la aplicación, y las tareas eran lo suficientemente variadas para evitar la saturación o aburrimiento de los niños.

Otra de las aplicaciones implementadas para el aprendizaje emocional es *CaptureMyEmotion*, destinada a niños del espectro autista, en el que se permite tomar fotos y videos al propio niño mientras registra simultáneamente

medidas fisiológicas (por ejemplo, el *arousal*⁷⁹), con el objetivo de favorecer el autorregistro. Puede que incluir información de tipo enciclopédica, con datos precisos e imágenes alusivas atravesadas por la propia experiencia, pueda permitir explotar la atención al detalle y el tipo de intereses de esta clase de niños, como por ejemplo su gusto por memorizar datos y cifras, y favorecer así el auto-registro.

Ya en el ámbito de la robótica y en lo que se denomina “tecnología de la afectividad”, Salvador, Silver y Mahoor (2015) emplearon el robot humanoide Zero en niños con y sin autismo. En esta propuesta los niños debían predecir el estado emocional del robot a partir de la información obtenida en sus expresiones faciales y, en una segunda fase, a partir de la expresión facial y la postura corporal del robot. No se encontraron diferencias en la predicción del estado emocional, pero la población con autismo tenía a cometer errores al identificar expresiones de miedo. La hipótesis de los autores era que los niños con autismo tendrían dificultades únicamente en el reconocimiento de emociones de lo que se llama “valencia negativa”, y que el lenguaje corporal podría ayudar en mayor medida que la sola expresión facial para reconocer las emociones. Para el modelado de los gestos, se empleó un programa de animación que incluía las descripciones del modelo de Ekman (1993) para la postura y el gesto corporal. Se solicitó a 50 actores que representaran las seis emociones básicas (tres representaciones por cada emoción), de las cuales se seleccionaron 18 imágenes al azar. En este experimento, se

⁷⁹ El *arousal* refiere a la medida de activación o respuesta del sistema nervioso central. Puede variar desde la sobre-activación en casos de emociones muy intensas, pasando por niveles óptimos de atención al entorno, hasta llegar a la infra-activación, como ocurre en el estado del sueño. (Pribram y McGuinness, 1975; Dorado y Fernández, 1997).

les decía a los niños que tenían que adivinar o predecir la emoción que el robot trataba de mostrar. Luego se los exponía a una serie de aproximaciones con el robot. Primero se les presentaba visualmente el robot con tarjetas que indicaban el nombre de una emoción y se les explicaba las reglas del juego. En un segundo momento, se exponía a los niños a mantener una conversación con el robot a través de un micrófono (los niños tenían acceso visual al robot, pero no estaban en la misma habitación). Luego de esta etapa de familiarización, los niños debían identificar las emociones que mostraba el robot y optar por una de ellas. Participaron aquí 22 niños de entre 7 y 13 años, de los cuales 11 tenían diagnóstico de autismo. Si bien no encontraron una diferencia estadísticamente significativa de rendimiento entre grupos de niños, la incorporación del gesto corporal como información extra generó diferencias de interpretación en determinadas emociones. Con la adición del gesto corporal, los niños con autismo tendían a elegir disgusto donde correspondía enojo y fallaban en predecir las verdaderas expresiones de disgusto. Los niños de desarrollo típico, en general, acertaban mejor con la información agregada del gesto corporal (excepto en los casos de sorpresa). En ambos grupos la expresión de felicidad resultó más ambigua al agregar el gesto corporal, lo que, según los autores, puede deberse a que la situación de modelado fue fabricada a partir de una historia que podía entremezclar ambas emociones⁸⁰. Un estudio similar, aunque con un robot menos expresivo a nivel facial, se utilizó para fomentar el reconocimiento de emociones a partir del lenguaje corporal en niños con autismo. Aquí también los niños tenían que estimar una emoción, a partir

⁸⁰ En la situación de modelado se les dijo a los actores que imaginaran en que se encontraban en la calle con un viejo amigo que hace mucho que no veían y que estaban muy contentos de verlo nuevamente.

de un conjunto de nueve emociones diferentes. En estos dos casos, diremos únicamente que la información contextual era limitada, ya que, aunque había mayor cantidad de detalle (por ejemplo, el gesto más humanoide que un pictograma o una fotografía), seguía siendo estática, fuera de un contexto o trama narrativa.

Respecto a las ventajas de emplear plataformas virtuales, se ha señalado en la literatura que ayudan a superar limitaciones y favorecen la autonomía, la atención, la incorporación de rutinas y el reconocimiento de tareas, la flexibilidad en el uso del espacio y el tiempo, el desarrollo educativo, la comunicación, el trabajo colaborativo y la independencia entre equipos; también se ha dicho que permiten el seguimiento del proceso de grupo, favorecen el acceso a la información y contenidos de aprendizaje, la superación de las limitaciones derivadas de las diferentes discapacidades, evitan la marginación estimulando mayores niveles de participación de los sujetos (Renilla Villalta *et al.*, 2011; Jiménez Lozano, Serrano Sánchez y Predes Espinoza, 2017). Otro aspecto significativo es que fomentan la flexibilidad cognitiva, ya que, en algunos casos, los usuarios pueden establecer sus propios itinerarios y tiempos, además de que es posible ajustar el tipo de *feedback* que se les brinda (Cuesta Gómez, 2012). Otros autores han señalado que el uso de plataformas virtuales tiene un impacto directo sobre la propia experiencia de los usuarios, debido a que estos muestran un aumento en la motivación para el aprendizaje y una mejora conductual (Orozco Cazco, Tejedor Tejedor y Calvo Álvarez, 2017; Pradilla, Belloso y Barboso, 2016; Calzadilla, 2002; Sánchez y Rodríguez Miranda, 2018). Por otro lado, en la medida en que dan respuesta a la diversidad de intereses y capacidades del alumnado y favorecen la inclusión educativa, así como también el desarrollo personal y social, las plataformas

virtuales resultan un recurso eficaz para trabajar la atención e impulsividad, al igual que el abordaje de contenidos específicos (Chousa Cortés, Martínez-Figueira y Raposo Rivas 2017).

Particularmente en el autismo, este tipo de recursos brinda más precisión sin requerir interacción social, es predominantemente visual y ayuda al aprendizaje social y la comunicación. También permite la presentación controlada de estímulos, ingresar tipo de información y claves personalizadas y crear ambientes que favorezcan el desarrollo de percepción-acción, así como otras habilidades de interacción y resolución de problemas.

Asimismo, como señalan Hourcade, Miller, Huebner y Liang (2013), el uso de dispositivos móviles puede ayudar a disminuir la ansiedad que en ocasiones puede llegar a expresar la población con autismo ante eventos inciertos e inesperados propios de las interacciones sociales⁸¹. Este tipo de interacciones resultan más predecibles y controladas, lo que causa una sensación de comodidad y genera mayores niveles de autoconfianza y enganche con la actividad.

Sin embargo, las críticas al uso de aplicaciones en general se han concentrado en que se debe dedicar tiempo y recursos de familiarización con su lenguaje y asegurar

⁸¹ Se ha señalado en la literatura una fuerte comorbilidad entre el síndrome de Asperger, trastornos del estado de ánimo y trastornos de ansiedad, por lo tanto, es legítimo cuestionar si esta reducción de la ansiedad responde a una reducción en los síntomas nucleares del autismo, o si se está interviniendo sobre otros cuadros psicopatológicos. Hay dos grandes hipótesis que explican la asociación entre síntomas de ansiedad y autismo: según la hipótesis del error social, el desajuste social de las personas con autismo tienden a exacerbar la ansiedad en esta clase de sujetos, y, según la hipótesis de la carga aleostática, la ansiedad es una respuesta ante el estrés crónico, fruto de la hiperactivación de las estructuras del sistema límbico (Pérez y Martos Pérez, 2009). En cualquier caso, es necesario abrir nuevos caminos de investigación que permitan una mayor discriminación diagnóstica, y desarrollar herramientas técnicas apropiadas que favorezcan el bienestar ante ciertas experiencias de sufrimiento.

las mismas oportunidades de acceso y un tipo de accesibilidad adecuada: competencia digital del alumnado e idoneidad del tipo de plataforma que utilizar (Orozco Cazco, Tejedor Tejedor y Calvo Álvarez, 2017). Otra de las críticas alude a que tienen escasa transferencia a la vida cotidiana y pueden generar una dependencia al tipo de estímulo.

Por otro lado, resulta difícil evaluar su nivel de transferencia a la vida cotidiana. En este sentido, Orozco Cazco, Tejedor Tejedor y Calvo Álvarez (2017) realizaron un metaanálisis para evaluar el efecto del uso de *software* educativo en alumnos de 6 a 12 años con necesidades educativas especiales. Encontraron que había una baja producción de estudios en esta área, así como también una supremacía de estudios en inglés sobre estudios anglohispanos (solo un 20 % en América Latina). La mayoría de las aplicaciones tenían una función instructiva (50 %), junto con un 40 % de aplicaciones con función evaluadora, y solo un 10 % tenía una función motivadora. Si bien la población con la que se realizaban los estudios era demasiado pequeña para sacar conclusiones (rango de 4 a 12 casos), encontraron que a nivel estadístico los resultados tenían un tamaño del efecto bajo⁸². Pese a ello, descubrieron que los alumnos de los grupos experimentales mejoraban un 72 % con el uso del software, mientras que el grupo de control, solo un 28 %. Las dificultades en los estudios de metaanálisis, como el de los autores, tiene que ver con las

⁸² El tamaño del efecto es una medida de estadística descriptiva, definida como el grado de generalidad que tienen los resultados obtenidos, es decir, en qué medida se espera encontrar estos mismos resultados encontrados en la población (Ledesma, Macbeth y Cortada de Kohan, 2008). Téngase en cuenta que, para un estudio experimental o cuasi experimental, se suele emplear muestras, ya que es imposible estudiar la totalidad de los casos posibles. En este caso, el tamaño del efecto, empleando la escala Cohen, es de 21, un valor bajo, por lo que indicaría un escaso nivel de generalidad para la población.

diferencias entre las técnicas de análisis y en el tipo de *software* utilizado, y con la variedad de asignaturas y diversidad de necesidades a las que apuntan.

Otros autores también han sugerido que ciertas plataformas deberían complementar a los tratamientos y/o al espacio aúlico. Se ha señalado que deberían implementarse diseños participativos, colaborativos e inclusivos, capaces de integrar el conocimiento de los expertos, los intereses de los niños y la experiencia de los diseñadores. Además, se ha sugerido que el diseño y el uso de aplicaciones deberían estar integrados en una experiencia lúdica, y que se debería considerar las contribuciones de los niños (véase, por ejemplo, Fletcher-Watson, Pain, Hammon, Humphry y McConachie, 2016).

Un ejemplo de diseño colaborativo es el de Constantin *et al.* (2017), que trabajaron con sujetos con CEA y retraso mental de entre 11 y 15 años. En aquella actividad, los usuarios tenían que interactuar con dos plataformas de tareas que involucraban habilidades de interacción social, una de ellas con recompensas por la ejecución de la tarea y la otra no, bajo el supuesto de que los niños con CEA parecen apreciar y preferir las recompensas. Los usuarios interactuaban con ambas plataformas y se iniciaba un proceso en el que se les preguntaba acerca de las recompensas y se los invitaba a modificarlas, o a incluirlas si no las tenía. Se llevó a cabo una tormenta de ideas (*brainstorming*) con los usuarios, en la que se puso a su disposición hojas, lápices e imágenes que pudieran inspirarlos. Cada niño podía identificar su hoja de ideas según el color de la tinta de la lapicera que habían empleado. Luego, se pidió a los niños que dieran más detalles acerca de sus ideas desde diferentes roles: como entrevistadores, siguiendo un guión de preguntas, como entrevistados que debían contestarlas o como camarógrafos. Los roles fueron intercambiándose

para que cada niño pudiera expresar sus ideas. En la tercera sesión se les solicitó que clasificaran las diferentes propuestas: a) opciones de recompensa dentro del sistema y b) recompensas asociadas al proceso. A continuación, se les pidió que eligieran una opción dentro del primer grupo para construir un prototipo. Para ello, se les dio hojas de papel con un marco de iPad, colores e imágenes de capturas de pantalla con diferentes elementos de la interfaz del sistema (menús, cajas, botones) para ir añadiendo a su diseño. Finalizada esta etapa, cada niño presentaba su diseño, mostrándolo y hablando sobre él con la misma dinámica de roles de entrevistador, entrevistado y operador de cámara empleada en la sesión anterior. Luego se les invitó a construir una recompensa (es decir, una señal de tarea realizada con éxito, por ejemplo, un cartel que dijera "Bien hecho"). Las recompensas podían ser de dos tipos: el mismo tipo de señal al finalizar cada fase, o señales según el rendimiento acumulativo. Así, a medida que acumulaban respuestas correctas, las recompensas podían incluir más elementos que la primeras veces (texto, animación, figurines, etc). Se les presentó los mismos materiales de diseño y dinámica de roles.

A partir del análisis de los datos, los investigadores encontraron que los niños tendían a elegir sus propios diseños y que preferían que el tipo de *feedback* fuera en aumento según su progreso. El tipo de *feedback* podía incluir texto, imagen, sonido y animación. Como el grupo con retraso mental tenía además otro tipo de diagnósticos, se replicó con nuevos grupos de niños para comparar los rendimientos entre los niños con CEA, con retraso mental y un tercer grupo con ambos diagnósticos. Se consultó con los maestros de los niños acerca del tipo de consignas del proyecto y se incorporaron las sugerencias recibidas (suprimir tipos de consignas que podían no ser

comprendidas, por ejemplo). El siguiente estudio tuvo el objetivo de identificar intereses (por ejemplo, animaciones electrónicas, tortugas ninja, *Pokémons*, desastres naturales, eventos históricos, datos fácticos de tipo enclopédicos, etc.). Los niños tenían que elegir sus preferidas dentro de las opciones sugeridas y anotar al final si tenían alguna otra preferencia. Luego se les presentó tarjetas con instancias específicas de dichas preferencias y tenían que clasificarlas entre las que les gustaban y las que no. Aquí también se les pidió que diseñaran una recompensa de proceso (según el rendimiento acumulado) según los diferentes niveles, incluyendo los mismos formatos que el estudio anterior (animación, sonido, fuegos artificiales y/o texto). En este trabajo, se incluían, en un marco de investigación de estudios previos sobre la problemática, los desafíos en la población con CEA y sus tipos de intereses, además de presentar como novedoso el trabajo y diseño colaborativo entre los usuarios, docentes y técnicos. Los autores también encontraron que los niños con CEA preferían en sus recompensas estímulos sociales (caras fijas de dibujos animados)⁸³ y mostraban preferencias por desarrollar y generar ideas acerca del tipo de recompensas por obtener. No hallaron diferencias significativas en el tipo de recompensas pre-dilectas en ambos grupos, y esto puede resultar relevante ya que muchos niños con CEA presentan, además, diagnóstico de retraso mental. Otro descubrimiento fue que los niños preferían las recompensas de proceso. En resumen, las ventajas de este estudio fueron que la aplicación reforzaba la motivación de los niños y que fue realizada

⁸³ Los autores mencionan que, pese a que aparece cierta dificultad para el reconocimiento de emociones, esta no se presenta en los casos de dibujos animados; este hallazgo es consistente con dichos antecedentes (Brosnan *et al.*, 2015, en Constantin *et al.*, 2017).

en un contexto de interacción cara a cara, con posibilidades de modificación de la aplicación según los propios intereses.

También Fletcher-Watson, Pain, Hammon, Humphry y McConachie (2016) emplearon un diseño colaborativo de una aplicación destinada a niños con CEA de 6 años de edad. A diferencia del estudio anterior, también se consultó con un grupo de personas que incluía a dos padres de niños con esta condición, dos profesores de niños en edad preescolar, un terapeuta del lenguaje y un adulto con CEA de alto funcionamiento. Se les preguntó si el proyecto les parecía relevante, si creían que podía tener un impacto a nivel terapéutico, y qué riesgos posibles consideraban necesario tener en cuenta considerando la población destinataria. El objetivo de la aplicación era brindar oportunidades de ensayar habilidades sociales en un contexto que fuera muy motivador para los niños. Se trabajó con niños que tenían, en su mayoría, diagnóstico de autismo y discapacidad intelectual. El diseño contó con varias etapas: la primera de ellas incluía actividades libres de dibujo y *collage* de varios tipos de imágenes. Luego se realizó un análisis del contenido de las producciones, considerando si las imágenes eran familiares o exóticas, si incluía animales, seres humanos u objetos, contenido apropiado al contexto (por ejemplo, un astronauta en la luna o un vaquero en la luna) y el tipo de diseño. En las entrevistas al grupo focal, surgió que la propuesta era interesante y atractiva, y que era deseable que incluyeran señales de tarea realizada con éxito de diferente tipo (texto, formas giratorias y vehículos), así como también la inclusión de indicaciones verbales, sonidos y animaciones. Asimismo, el grupo focal expresó preocupación por la generalización de estas habilidades a la vida real y el riesgo de que los niños se quedaran fijados en algún estímulo que fuera de

su agrado y buscaran repetirlo constantemente en lugar de jugar con la aplicación. La siguiente fase fue el desarrollo de la aplicación siguiendo diferentes niveles de dificultad. Siguiendo la tradición de la psicología experimental, el juego tenía objetivos y distractores. La *app* fomentaba dos habilidades prosociales: atender a otras personas y seguir señales sociales. En la primera parte del juego, los niños debían identificar un personaje humano (objetivo), y no otra clase de elementos presentes en la pantalla (distractores). En la segunda parte del juego, el personaje miraba o señalaba un objeto que podía ser localizado por el jugador para pasar a la siguiente pantalla. Objetivos y distractores provinieron de los diseños de los niños, y las pantallas de inicio y transición tuvieron estímulos neutros (imagen estática sin audios) para evitar una fijación obsesiva de los niños en repetir dicho estímulo. Las recompensas aparecían cuando el niño señalaba los objetivos, y eran de tres tipos: recompensas geométricas, recompensa social (por ejemplo, acróbatas que saltaban y realizaban aplausos en la escena) y recompensa de objeto de interés (por ejemplo, que apareciera un tren, generalmente de interés para los niños). No se utilizaron recompensas negativas ante respuestas incorrectas. Si bien no era posible hacer un sustrato narrativo porque el juego era muy simple, una de las tareas implicaba ayudar a una persona en un supermercado, como una forma básica de estructura narrativa. Las limitaciones de este estudio fueron que no se incluyeron tantas características personalizables como se hubiera deseado. Los autores también señalan que, para comprobar si el diseño colaborativo es más eficaz, debería compararse el rendimiento en aplicaciones con y sin diseño colaborativo.

6. ¿Pueden las plataformas virtuales contribuir a la experiencia y favorecer la lectura de expresiones faciales, el reconocimiento de emociones y las claves sociales?

Es de vital importancia tener en cuenta las diferencias individuales entre las personas y las necesidades particulares, y generar los apoyos necesarios para favorecer el desarrollo, porque, como dicen Baña y Losada-Puente (2017, p. 131), “si una persona presenta alteraciones en el desarrollo, tiene dificultades para comprender el mundo en el que vive, se le reducen las oportunidades para elegir, tomar decisiones y entrenar otros procesos mentales, difícilmente puede adquirirlos o ampliarlos”. Para pensar, entonces, el diseño de plataformas virtuales, sería necesario distinguir, en primer lugar, los fines que persigue. Es decir, si pretende: a) evaluar la presencia o ausencia de una habilidad, b) enseñar una habilidad ausente, o c) afianzar una habilidad presente, pero de ejecución deficitaria. La ejecución puede ser deficitaria debido a la cantidad de recursos que emplea el sujeto –excesiva o insuficiente–, la calidad de la estrategia –ajustada o desajustada– o de qué modo se evaluó el contexto de aplicación, es decir, si el sujeto pudo diferenciar los contextos de uso y establecer las variaciones correspondientes⁸⁴. También se debería tener en cuenta el contenido, el tipo de diseño y la cantidad de información que aportará. En cuanto al contenido, si se trata de tareas de reconocimiento de emociones, por ejemplo, se debe considerar si se emplean pictogramas o videos, o fotos de personas conocidas para los sujetos. El tipo de

⁸⁴ Pensemos en habilidades sociales: por ejemplo, la capacidad de hacernos escuchar en un contexto lleno de ruido, donde debemos elevar la voz, o de hacernos escuchar en un contexto de examen o en un velatorio, donde debemos casi susurrar, o en las diferencias en el tipo de contacto físico que tenemos con alguien familiar por contraposición a un desconocido, etc.

diseño puede ser por modelado, donde los sujetos tienen que ejecutar los diferentes elementos que componen una habilidad, o por aprendizaje asociativo, o por colaboración participativa para el tipo de situaciones que se abordan, etc. Respecto a la cantidad y tipo de información que brindará, nos referimos a si la información dada es estática, molecular o contextual, incluyendo descripciones de una situación o trama narrativa, información ampliada deco-dificada en lenguaje corporal o información dinámica que involucre reconocimiento de emociones en tiempo real.

Lo que intentamos defender aquí es que, si bien el aprendizaje de tipo socioemocional es ecológico, corporizado, atravesado por la experiencia vivida, así como por su contexto de emergencia y adquisición, y el uso de plataformas virtuales tiene sus limitaciones, estas podrían servir de apoyo a la experiencia en casos particulares, como en el autismo. También es necesario tener en cuenta que el empleo de medios digitales no asegura ni reemplaza la competencia que se intenta fortalecer. A su vez, si partimos de que el desarrollo en el autismo es diferente a lo esperado en los casos de desarrollo típico, no podemos pretender que tengan el mismo tipo de funcionamiento social, pero sí podemos aspirar a reducir el malestar asociado a situaciones sociales, y a incrementar las habilidades socioemocionales necesarias para la vida diaria, como la comunicación y la resolución conjunta de problemas.

Creemos que el uso de plataformas virtuales, pese a sus limitaciones y a todos los desafíos que plantea en la actualidad, permitiría: 1)hacer explícitos los razonamientos subyacentes a ciertas prácticas sociales que involucran la expresión de estados emocionales; 2) reconocer estímulos socialmente relevantes, entre ellos, gestos y expresiones faciales; 3) analizar los intentos fallidos, pudiendo revertir y ensayar posibles respuestas sin la velocidad de

las interacciones en vivo (que, dado la gran cantidad de información contextual implícita y explícita puede en un principio resultar abrumador para sujetos con esta condición), y 4) afianzar el rendimiento socialmente competente mediante el conocimiento explícito de patrones de comportamiento apropiados y entrenamiento en la lectura de emociones a partir de las expresiones faciales y el lenguaje corporal.

En este sentido, las nuevas tecnologías permiten ajustarse a la experiencia de cada usuario y tener la posibilidad de revertir una secuencia de operaciones, recortar información para su mejor procesamiento y experimentar una temporalidad diferente a las situaciones sociales cara a cara, teniendo un margen mayor para seleccionar la respuesta adecuada (Sanroma Giménez, Lázaro Cantabrana, Gisbert Cervera, 2017).

Las posibilidades de conocimiento y acceso a la información para la vida y la independencia en calidad de vida pasan por las oportunidades y posibilidades de experimentar, aprender y participar de los instrumentos que la propia sociedad ha ido generando para ello (Baña y Losada-Puente, 2017, p. 133).

Además, el auge cada vez mayor de las nuevas tecnologías en esta población se debe al hecho de que las personas con TEA sienten una afinidad especial por estas, por lo cual resulta una vía óptima para el entrenamiento en áreas que presentan desafíos (Pérez-Fuster, 2017). También se ha señalado que:

La singularidad de cada ser humano requiere un estilo de acercamiento y desarrollo integral, en atención a su experiencia vital y al modo peculiar de descubrir las óptimas interacciones entre todas las personas del ambiente de aprendizaje, construido en una fecunda y activa colaboración (Medina Rivilla y Rodríguez Serna, 2016, p. 2).

Las personas con CEA muchas veces tienen predilección por medios electrónicos, fuertes habilidades de procesamiento visual y motivación hacia las plataformas virtuales (Cuesta Gómez, 2012). Es por estos motivos por lo que las plataformas virtuales pueden ayudar a decodificar señales que se han visto obstaculizadas en el desarrollo temprano, pero no pueden desprenderse del lugar de apoyo complementario, ya que es la experiencia situada, directa, embebida de relaciones interpersonales la que favorece una lectura “*online*” de los cuerpos y, por lo tanto, un acceso a la lectura de mentes de los demás.

Teniendo en cuenta todas estas consideraciones, creamos que el diseño de una aplicación que permitiera enriquecer la experiencia debería ser elaborada mediante un diseño colaborativo, aplicado en el marco de una experiencia de taller o de grupos colaborativos.

Una aplicación de este tipo también debería contar con una selección de estímulos significativos para la lectura de claves sociales, como, por ejemplo, situaciones sociales versus situaciones no sociales, componentes en formato verbal y no verbal o paraverbales (por ejemplo, el tono de voz), movimientos y estructuras espaciales, escenarios de comunicación, atribución de contenido a gestos y expresiones faciales, etc. Además, la aplicación debería tener algún tipo de *feedback* sobre el rendimiento del usuario que resulte llamativo y atractivo para la población a la que está destinada.

En la literatura se ha señalado también que cualquier plataforma virtual debe atender a una diversidad de aspectos, tales como la integración de TIC al aula, el fomento para la adquisición de equipos específicos y ayudas técnicas de acceso al ordenador, la formación del profesorado en *software* educativo, criterios de diseño universal, pautas de accesibilidad en el diseño web, investigación sobre

metodología y didáctica de aplicación de TIC en población con necesidades educativas especiales, colaboración entre los profesionales y la comunidad educativa, sensibilización social sobre la importancia de la accesibilidad en estas poblaciones (Luque Rodríguez, 2009, en Orozco Cazco, Tejedor Tejedor y Calvo Álvarez, 2017), además de facilitar la presencia y el acceso a las TIC, revisar la adaptabilidad y accesibilidad y la alfabetización digital (Orozco Cazco, Tejedor Tejedor y Calvo Álvarez, 2017)

También, como afirman Orozco Cazco, Tejedor Tejedor y Calvo Álvarez (2017), resulta interesante pensar a qué tipo de necesidad están apuntando las aplicaciones que diseñamos, cuáles son las funciones de *software* aplicadas a estos campos y qué tamaño del efecto tiene su implementación. En este sentido, es una muy buena pregunta orientadora cómo es posible generar oportunidades de aprendizaje basado en la experiencia, ya que lo importante es ofrecer experiencias significativas, donde algunas de las actividades pueden ser enriquecidas con instrumentos digitales, pero adoptando una estrategia de diseño colaborativo para que la sola construcción de la herramienta sea una experiencia interactiva en sí misma.

Referencias bibliográficas

- American Psychiatric Association y American Psychiatric Association (Eds.). (2013). *DSM-V* (5.^º ed). Washington, D.C: American Psychiatric Association.
- Bañá, M. y Losada-Puente, L. (2017). “La intervención en el desarrollo de las personas con Trastornos del Espectro del Autismo y las Tecnologías de la Información y la Comunicación”. En S. Pérez-Aldeguer, G.

- Castellano-Pérez y A. Pina-Calafi (Eds.), *Propuestas de Innovación Educativa en la Sociedad de la Información* (pp. 122-135). Eindhoven, NL: Adaya Press.
- Baron-Cohen, S., Lombardo M. V., Auyeung, B., Ashwin, E., Chakrabarti, B. y Knickmeyer, R. (2011). "Why Are Autism Spectrum Conditions More Prevalent in Males?". *Plos Biology*, 9(6), pp. 1-10.
- Bhat, A. N., Galloway, J. C. y Landa, R.J. (2010). "Social and Non-Social Visual Attention Patterns and Associative Learning in Infants at Risk for Autism". *Journal of Child Psychology and Psychiatry*, 51 (9), pp. 989-997.
- Burdman, F. G. (2016). "El post-cognitivismo en cuestión: extensión, corporización y enactivismo". *Principia: An International Journal of Epistemology*, 19 (3), pp. 475-495
- Calzadilla, M. E. (2002). "Aprendizaje colaborativo y tecnologías de la información y la comunicación". *Revista Iberoamericana de Educación*, 29(1), pp. 1-10. Recuperado de <https://bit.ly/2OmLjW1>.
- Castillo Aguilar, T.A., Pérez de Céliz Herrero, C., Muñoz, C., Somodevilla García, M., Pineda Torres, I., De Alba Aquilar, K. y Romero Romero, E. (2016). *AUTHIC: Herramienta computacional para niños con espectro autista. XVIII Simposio Internacional de Informática Educativa* (pp. 409-414). Puebla: Ediciones Universidad de Salamanca.
- Constantin, A., Johnson, H., Smith, E. y Lengyel, D. (2017). "Designing computer-based rewards with and for children with Autism Spectrum Disorder and or Intellectually Disability". *Computers in Human Behavior*, 75, pp. 404-414.

- Cuesta Gómez, J. (2012). "Tecnologías de la información y la comunicación: aplicaciones en el ámbito de los trastornos del espectro del autismo". *Revista Siglo Cero*, 43 (2), pp. 6-25.
- Crisol Moya, E. y Domingo Martos, L. (2015). "Aplicación multimedia para trabajar las emociones con alumnado con síndrome de Asperger". *Revista científica electrónica de Educación y Comunicación en la Sociedad del Conocimiento*. 16(1), pp. 1-26. Recuperado de <https://bit.ly/2Ub7dCr>.
- Dorado, M. A. y Fernández, I. (1997). "Psychophysiological and self-reported arousal: Differences between Type A and Type B behavior pattern subjects". *Revista de Psicología Social*, 12(1), pp. 43-58.
- Droucker, D., Curtin, S. y Voloumanos, A. (2013). "Linking infant-directed speech and face preferences to language outcomes in infants at risk for Autism Spectrum Disorder". *Journal of Speech, Language, and Hearing Research*, 56 (2), pp. 567-576.
- Del Moral Pérez, M. E., Bellver Moreno, M. C. y Guzmán Duque, A. P. (2018). "Instrumento para evaluar la potencialidad creativa de app orientadas al diseño de relatos digitales personales". *Digital Educational Review*. 33, pp. 284-305. Recuperado de <http://greav.ub.edu/der/>.
- Chousa Cortés, C., Martínez-Figueira, M. E. y Raposo Rivas, M. (2017). "Las TIC para la intervención educativa en TDAH: un estudio bibliométrico". *Perspectiva Eduacional*, 56 (3), pp. 142-161.
- Ekman, P. (1993). "Facial expressions and emotion". *American Psychological Association*, 48 (4), pp. 384-392.

- Enticott, P., Kennedy, H. A., Johnston, P. J., Rinehart, N. J., Tonge, B. J., Taffe, J. R. y Fitzgerald, P. B. (2014). "Emotion recognition of static and dynamic faces in Autism Spectrum Disorder. *Cognition and Emotion*, 28(6), pp. 1110-1118.
- Falco, M. (2017). "Reconsiderando las prácticas educativas: TICs en el proceso de enseñanza-aprendizaje". *Tendencias Pedagógicas*, 1(29), pp. 59-76.
- Fletcher-Watson, S., Pain, H., Hammond, S., Humphry, A. y McConachie, H. (2016). "Designing for young children with Autism Spectrum Disorder: A case study of an iPad app". *International Journal of Child-Computer Interaction*, 7, pp. 1-14. Recuperado de <https://bit.ly/2HRAeME>.
- García Arnanz, L., Herráez Aón, M., Olivares Romero, E., Sánchez, M. C. S. y Andrés-Roqueta, C. (2018). "Evaluación de la competencia emocional en niñas/os con Trastorno del Espectro Autista a través de una APP novedosa". *Àgora de Salut*, 5, pp. 217-226.
- García-Guillén, S., Garrote Rojas, D. y Jiménez Fernández, S. (2016). "Uso de las TIC en el Trastorno de Espectro Autista". *Edmetic*, 5(2), pp. 134-157.
- García Alonso Navarrete, M. A. (2018). "Tecnologías de la Información y Comunicación para niños con Trastorno del Espectro Autista". *Inventio*, 14(32), pp. 5-8.
- Greenspan, S. y Wieder, S. (2006). *Engaging autism. Covering all types of ASD, including Asperger's Syndrome, from the earliest signs in infancy through the school years and on into adulthood*. Philadelphia: First Da Capo Press.
- Guzmán, G., Putrino, N., Martínez, F. y Quiroz, N. (2017). "Nuevas tecnologías: puentes de comunicación en el Trastorno del Espectro Autista (TEA)". *Terapia Psicológica*, 35(3), pp. 247-258.

- Hourcade, J. P., Williams, S. R., Miller, E. A., Huebner, K. E., y Liang, L. J. (2013). "Evaluation of tablet apps to encourage social interaction in children with Autism Spectrum Disorders". En Mackay, W. (Presidencia) Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI'13 (p. 3197). París.
- Jiménez Losano, M. D., Serrano Sánchez, J. L. y Prendes Espinosa, M.P. (2017). "Estudio de caso de la influencia del aprendizaje electrónico móvil en el desarrollo de la comunicación y el lenguaje con un niño con TEA". *Educar*, 53(2), pp. 419-443.
- Kirby, R., Forlizzi, J. y Simmons, R. (2010). "Affective social robots". *Robotics and Autonomous Systems*, 58(3), pp. 322-332.
- Ledesma, R., Macbeth, G. y de Cohan, N.C. (2008). Tamaño del efecto: revisión teórica y aplicaciones con el sistema estadístico. *Vista. Revista Latinoamericana de Psicología*. 40(3), pp. 425-439.
- Malinverni, L., Mora-Guiard, J.. Padillo, V.. Valero, L., Hervás, A, y Pares, N. (2017). "An inclusive design approach for developing videogames for children with Autism Spectrum Disorder". *Computers in Human Behavior*, 71, pp. 535-549.
- Medina Rivilla, A. y Rodríguez Serna, C. (2016). "Potenciar las capacidades de las personas. Modelo para facilitar la comunicación con estudiantes del espectro autista". *Revista Nacional e Internacional de Educación Inclusiva*, 9(1), pp. 1-12.
- Montañez Ruiz, J.J. (2017). *Las TIC como contribución al desarrollo lectoescritor del alumnado con TEA*. Tesis de maestría de la Universidad de Jaén. Dependencia e Igualdad en la Autonomía Personal. Recuperado de <https://bit.ly/2Vc7T6U>.

- Montagut Asunción, M., D'Ocon Giménez, A. y Pastor Cerezuela, G. (2017). "Interacción social y comunicación temprana en hermanos de niños con diagnóstico de Trastorno del Espectro Autista: una revisión de la literatura". *International Journal of Developmental and Educational Psychology. Revista INFAD de Psicología.*, 4(1), pp. 373-388.
- Orozco Cazco, G. H., Tejedor Tejedor, F. J. y Calvo Alvarez, M. I. (2017). "Meta-análisis sobre el efecto de Software educativo en alumnos con Necesidades Educativas Especiales". *Revista de Investigación Educativa*, 35(1), pp. 35-52.
- Osterling, J. A., Dawson, G., Munson, J. A. (2002). "Early recognition of 1-year-old infants with Autism Spectrum Disorder versus mental retardation". *Developmental Psychopathology*, 14, pp. 239-251.
- Peretti, L. (2012). "La cognición social corporizada según Goldman y de Vignemont: algunas incompatibilidades con el enfoque de S. Gallagher". En Salvático, L., Bozzoli, M., Pesenti, L. (Eds.), *Epistemología e historia de la ciencia* (pp. 447-454). Córdoba, Argentina: Editorial UNC.
- Pérez-Fuster, P. (2017). *Enhancing skills in individuals with autism spectrum disorder through technology-mediated interventions*. Tesis doctoral de la Universidad de Valencia (España). Recuperado de <https://bit.ly/2I3WNxd>.
- Pérez, D. y Lawler, D. (comp). (2017). *La segunda persona y las emociones*. Ciudad Autónoma de Buenos Aires: SADAFA.
- Pérez, P. I. y Martos, J. M. P. (2009). "Síndrome de Asperger y autismo de alto funcionamiento: comorbilidad con trastornos de ansiedad y del estado de ánimo". *Revista de Neurología*, 48(2), pp. 31-34.

- Perrault, A., Chabi, L., Bigouret, F., Oppetit, A., Cohen, D., Plaza, M. y Xavier, J. (2018). "Comprehension of conventional gestures in typical children, children with Autism Spectrum Disorders and children with language disorders". *Neuropsychiatrie de l'Enfance et de l'Adolescence. ScienceDirect*. Disponible en <https://bit.ly/2TP2K32>.
- Pradilla, J., Belloso, O. y Barbosa, J. (2016). "Factores que determinan la efectividad de las TIC en la mediación tecnológica del aprendizaje". *Revista electrónica de Humanidades, Educación y Comunicación Social*. 11(22), pp. 243-259.
- Pribram, K. H., y McGuinness, D. (1975). "Arousal, activation, and effort in the control of attention". *Psychological Review*, 82(2), pp. 116-149.
- Ramírez, V. y Cassinero, S. (2014). Realidad Aumentada -Trabajo Cooperativo; Nivel Inicial. *Congreso Iberoamericano de Ciencia, Tecnología, Innovación y Educación*. Artículo 1669. Recuperado de <https://bit.ly/2OBRQNj>.
- Renilla Villalta, M. R., Pedrero Esteban, D. A., Sánchez Cabaco, D. A., Estévez Villa, D. J. (2011). "Evolución tecnológica en la enseñanza virtual de las personas con TEA". *International Journal of Developmental and Educational Psychology*, 4(1), pp. 383-392.
- Salvador, M., Silver, S. y Mahoor, M. (2015). "An Emotion Recognition Comparative Study of Autistic and Typically-Developing Children Using the Zeno Robot". 2015 IEEE International Conference on Robotics and Automation (ICRA). Washington State Convention Center. Seattle, Washington, mayo 26-30.

- Sanabria, F. J. (2017). "Efecto de la experiencia corporal de frío sobre la calidez interpersonal hacia grupos sociales. Un estudio desde la cognición corporizada". *Wímbalu, Revista de Estudiantes de la Escuela de Psicología*, 12(1), pp. 57-67.
- Sanromà-Giménez, M., Lázaro-Cantabrana, J.L., Gisbert-Cervera, M. (2017). "La tecnología móvil. Una herramienta para la mejora de la inclusión digital de las personas con TEA". *Psicología, Conocimiento y Sociedad* 7(2), pp. 227-251.
- Segovia Cuéllar, A. (2017). "El papel de la acción y la socialización afectiva en el desarrollo del entendimiento social en humanos". En Gamboa, L., González, J., Hüni, H., González, M., Moncada, J., Alloa, E., Franco, L. (2017). *Pensar el cuerpo* (Gamboa L. y González J., Eds.) (pp. 357-374). Barranquilla, Colombia: Editorial Universidad del Norte.
- Serrano Sánchez, J. L. y Prendes Espinosa, M.P. (2015). *Estudios sobre educación*, 28, pp. 187-210.
- Silva Sánchez, G. y Rodríguez Miranda, F. de P. (2018). "Una mirada hacia las TIC en la educación de las personas con discapacidad y con Trastorno del Espectro Autista: Análisis temático y bibliográfico". *EDMETIC, Revisita de Educación Mediática y TIC*, 7(1), pp. 43-65.
- Suria Martínez, R. (2012). "Redes sociales online y su utilización para mejorar las habilidades sociales en jóvenes con discapacidad". *Escritos de Psicología*, 5(3), pp. 16-23.
- Terrazas Acedo, M., Sánchez Herrera y Becerra Traver (2016). "Las TIC como herramienta de apoyo para personas con Trastorno del Espectro Autista". *Revista Nacional e Internacional de Educación Inclusiva*. 9(2), pp. 102-136.

- Trevarthen, C. (1998). "The concept and foundations of infant intersubjectivity". En S. Bråten (Ed.), *Intersubjective Communication and Emotion in Early Ontogeny*, (pp. 15-46). Cambridge: Cambridge University Press.
- Tronick, E. Z., Als, H. y Brazelton, T.B. (1978). "The infant's response to entrapment between contradictory messages". *Journal of the American Academy of Child Psychiatry*, 17, pp. 1-13. www.panaacea.org, visitado el 11 de diciembre de 2018.

Gestos y palabras

*Los gestos de maestras y alumnos durante la enseñanza
de vocabulario⁸⁵*

ALEJANDRA MENTI (UNC – CONICET) Y CELIA ROSENBERG (CONICET – CIIPME)

1. Introducción

Este trabajo tiene por objeto analizar los gestos realizados por docentes y alumnos en el proceso de construcción de significados en los intercambios de enseñanza y aprendizaje. Específicamente, interesa analizar aquellas secuencias de habla en las que las maestras o los alumnos se centran en un vocabulario específico, no familiar o poco frecuente, ya sea para explicar o aclarar su significado, para corregir el uso de un término o para reparar una ruptura en la comunicación en el aula.

Desde una perspectiva psicolingüística, sociopragmática (Tomasello, 2003) y experiencial (Nelson, 2007), se postula que el aprendizaje lingüístico y conceptual infantil se produce en el marco de los contextos de interacción

⁸⁵ Publicado originalmente en Íkala. Revista de Lenguaje y Cultura, 22(3), 455-475. DOI: 10.17533/udea.ikala.v22n03a06. Este artículo se financió con aportes del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) y de la Secretaría de Ciencia y Tecnología (SECyT).

en los que participan los niños. En esta perspectiva, diversas investigaciones pusieron de manifiesto que, para poder inferir y comprender el significado de las palabras desconocidas, los niños se apoyan en el sostén social, pragmático, lingüístico y paralingüístico que les proporcionan sus interlocutores (Beals y Tabors, 1995; Weizman y Snow, 2001).

En consonancia con este encuadre teórico, diversos estudios desarrollados en el marco del análisis de la conversación y de la sociolingüística interaccional mostraron que el significado de las interacciones se construye en la yuxtaposición de las expresiones lingüísticas con otros sistemas semióticos, como los gestos, la dirección de la mirada, la orientación corporal, las expresiones faciales, los movimientos de todo el cuerpo, la postura, la variación en la entonación y la particular estructuración del contexto de actividad, así como los artefactos involucrados en la actividad (Lemke, 1998; Goodwin, 2000, 2007; Mondada, 2001; Barowy y Smith, 2008).

En las últimas décadas, una serie importante de investigaciones enmarcadas en diversas líneas teóricas centraron su objeto de estudio en el análisis sistemático de los gestos y, principalmente, en el rol que estos desempeñan en las interacciones (Kendon, 1980; McNeill, 1985, 1992, 2005; Goodwin y Heritage, 1990; Levinson, 1997; Wells, 1998; Goodwin, 2000, 2007; Orellana, Martínez, Lee y Montano, 2012, entre otros).

A pesar de que algunos de estos estudios han definido a los gestos como elementos separados del lenguaje, desde la perspectiva prevalente en otros trabajos (McNeill, 1992, 2005, 2016; Kendon, 2000) se ha postulado que los gestos generan un sistema integrado con el lenguaje de las personas. Asimismo, estos estudios pusieron de manifiesto el

modo en el que los pensamientos son transformados por medio de la gestualidad durante la comunicación (Goldin-Meadow, 2014).

Es importante resaltar que esta corriente de investigación se destaca y diferencia claramente de otras al sostener que las expresiones lingüísticas y los gestos constituyen modos alternativos para expresar un mismo significado (McNeill, 1985). Mientras que el habla es lineal, segmentada y compuesta por unidades más pequeñas, el gesto, por oposición, es global y sintético; en tal sentido, un solo gesto puede manifestar un conjunto de significados. Los gestos resultan en particular importantes cuando se presentan lagunas lingüísticas, debido a que constituyen una modalidad alternativa que expresa visualmente características del concepto aludido. De este modo, le permiten al interlocutor continuar con su discurso y darse a entender, aun cuando este no disponga de las palabras específicas que requiere en ese momento (McNeill, 1992; Wells, 1998; Lemke, 1999; Roth, 2000; Krauss, Chen y Gottesman, 2000).

En el marco de la psicología cognitiva, McNeill (1985, 1992, 2005, 2016) definió a los gestos como aquellos movimientos de las manos y de los brazos que acompañan el habla. Son la imagen intrínseca del habla (McNeill, 2016). Asimismo, clasificó a los gestos en cuatro tipos básicos: gestos de orquestación o rítmicos, deícticos, icónicos y metafóricos.

Los gestos de orquestación son los movimientos de la mano, simples y rápidos, que durante los intercambios conversacionales acompañan a las palabras más importantes del enunciado y, por tanto, cumplen la función de enfatizar las relaciones en el discurso. Asimismo, los gestos de orquestación funcionan como gestos interactivos, que sirven para regular la coordinación de los turnos de habla, para buscar o solicitar una respuesta, o para corroborar la

comprensión por parte de la otra persona. Por su parte, los *gestos deícticos* son aquellos que el interlocutor emplea ya sea para señalar entidades, fenómenos o cualidades concretas, o para reemplazar el habla. Los *gestos icónicos* categorizan a los movimientos de la mano-brazo que mantienen una relación perceptual formal de similitud con el contenido semántico de la expresión lingüística a la que acompañan. Por último, los *gestos metafóricos* se asemejan a los gestos icónicos y deícticos en que ellos representan o señalan una imagen visual; sin embargo, las imágenes a las cuales hacen referencia son conceptos abstractos. El contenido abstracto puesto de manifiesto por medio de este último tipo de gestos se relaciona no solo con las imágenes vinculadas a los objetos, espacios, movimientos, etc., sino también con el conocimiento lingüístico y cultural (McNeill, 1985; Roth, 2001).

En el ámbito educativo, algunas investigaciones atendieron al papel que cumplen los gestos durante la enseñanza y el aprendizaje. Así, por ejemplo, Roth (2001) identificó dos posibles funciones generales de los gestos en el aula: por un lado, el docente puede valerse de la información proporcionada por los gestos que realizan sus alumnos para evaluar si han comprendido o no el tópico de la clase; por otro, los gestos del docente pueden servir de apoyo o recurso adicional para que los alumnos comprendan el discurso del maestro. Esto se debe a que los gestos, al transmitir información en formato viso-espacial de modo simultáneo y sincronizado con las expresiones verbales (McNeill, 1992), enriquecen el mensaje transmitido en el discurso. Por otra parte, Mehrabian (1967) observó que los docentes se valen de expresiones verbales acompañadas de gestos con el propósito de disminuir la distancia

interaccional que existe entre ellos y los alumnos. De esa manera, generan inmediatez, es decir, un contexto de cercanía y proximidad física y psicológica con los alumnos.

En consonancia con este planteo, en diversos trabajos se ha señalado que las instrucciones orales acompañadas de gestos promueven mejor el aprendizaje que las mismas instrucciones transmitidas sin gestos (Gaythwaite, 2005; Singer y Goldin-Meadow, 2005; Migdalek y Rosemberg, 2012). En este sentido, Babad, Avni-Babad y Rosenthal (2004) destacaron que, cuando el docente conoce la importancia que adquieren los gestos en las interacciones del aula y los emplea de manera consciente, los alumnos pueden aprender mejor el contenido de la clase. En esta misma línea, Migdalek y Rosemberg (2012) observaron que, durante la planificación de instancias de juego en el jardín de infantes, las maestras emiten gestos cuando se refieren a aspectos centrales de la consigna. De este modo, los gestos podrían orientar la comprensión por parte de los niños de referentes abstractos, palabras desconocidas y reglas del juego.

Muchos de los trabajos que analizaron el papel de los gestos en las conversaciones que tienen lugar en la escuela primaria se focalizaron en clases de matemática (Goldin-Meadow, Kim y Singer, 1999; Goldin-Meadow, 2003, 2006; Goldin-Meadow y Singer 2003; Singer y Goldin-Meadow, 2005; Cook, Mitchell y Goldin-Meadow, 2008, entre muchos otros). Así, por ejemplo, el estudio experimental de Singer y Goldin-Meadow (2005) puso de manifiesto que los gestos realizados por las docentes, que transmiten información distinta a la expresada oralmente, ofrecen a los alumnos un camino alternativo para resolver el problema en cuestión y, en consecuencia, pueden facilitar el aprendizaje de matemática.

Por su parte, Cook, Mitchell y Goldin-Meadow (2008) observaron que los niños que emitieron gestos durante el aprendizaje de un nuevo concepto, con el tiempo, conservaron el conocimiento adquirido. Por el contrario, los niños a los que se solicitó solamente hablar -y no gestualizar- durante el aprendizaje del concepto matemático nuevo no consolidaron su aprendizaje a largo plazo. A partir de estos resultados, los investigadores llegaron a la conclusión de que los gestos pueden cumplir un rol importante en el aprendizaje, tal vez porque proporcionan a los alumnos, como alternativa, una manera encarnada de representar nuevos conceptos.

Son menos numerosas las investigaciones que analizaron el papel de los gestos durante clases de ciencias en la escuela primaria y secundaria (Crowder y Newman, 1993; Lemke, 1998; Newman y Crowder, 1991; Roth, 2000, 2002; Roth y Welzel, 2001; Roth y Lawless, 2002; Menti y Alam, 2014) y en el nivel terciario y universitario (Lemke, 1999; Furuyama, 2000; Babad, Avni-Babad y Rosenthal, 2004; Edwards, 2009; Kelly, Özyürek y Maris, 2010). Entre los trabajos que analizaron la función de los gestos de los alumnos durante las clases de ciencias en la escuela primaria se destacan las investigaciones de Newman y Crowder (Crowder y Newman, 1993; Newman y Crowder, 1991). Sus resultados pusieron de manifiesto que, cuando los alumnos tienen que explicar un conocimiento nuevo, en general no son verbalmente explícitos. Los gestos constituyen un modo de expresión alternativo -o el único modo- a través del cual pueden explicar la comprensión del conocimiento nuevo (Newman y Crowder, 1991). Asimismo, observaron que, durante las explicaciones, los alumnos emplean el lenguaje y los gestos como recursos comunicativos disponibles. La mayoría de los gestos, en general, redundan o amplían la información proporcionada mediante las

expresiones lingüísticas, o son empleados de forma alternativa como portadores de significado científico. En este sentido, encontraron que los gestos ayudan a que los alumnos expliquen con claridad sus conocimientos científicos, especialmente cuando no recuerdan el término científico preciso. Los alumnos se valen de gestos para ampliar sus explicaciones verbales y producir, así, un todo coherente.

Por su parte, Roth y colaboradores analizaron el papel que cumplen los gestos de los alumnos a lo largo del discurso durante clases de ciencias de la escuela secundaria (Roth, 2000, 2002; Roth y Welzel, 2001; Roth y Lawless, 2002, entre otros). En sus trabajos, Roth observó que los alumnos se valen de gestos icónicos y deícticos cuando no poseen el vocabulario científico apropiado para explicar y describir los conceptos científicos. Sin embargo, a medida que incrementan el dominio del discurso científico, los gestos ya no preceden al habla, sino que comienzan a producirse en simultáneo (Roth, 2000). Asimismo, Roth y Welzel (Roth y Welzel, 2001; Roth, 2002) pusieron de manifiesto que los gestos metafóricos facilitan a los alumnos el aprendizaje de conceptos abstractos, en presencia de los objetos materiales. El campo perceptual y los gestos cumplen una importante función de andamiaje en el desarrollo del lenguaje científico de los estudiantes, ya que asumen una función representativa (Roth, 2000, 2002; Roth y Welzel, 2001; Roth y Lawless, 2002, entre otros).

Por su parte, Menti y Alam (2014) analizaron los gestos que hicieron las maestras durante la enseñanza de palabras desconocidas o poco familiares a los alumnos de primaria. A diferencia de lo señalado en trabajos anteriores, que relacionan el uso de gestos con las palabras empleadas por el hablante (Singer y Goldin-Meadow, 2005), las autoras observaron que los gestos de las docentes no siempre respondían al tratamiento de una palabra desconocida,

sino a aspectos relacionados con los conocimientos previos de los alumnos. En efecto, cuando se mencionaba una palabra poco familiar a los niños, las maestras no recurrían en todos los casos a información gestual para acompañar el término. Cuando la palabra desconocida o poco familiar era pronunciada por la maestra en el marco de un tópico conocido por los alumnos, no yuxtaponía gestos a su emisión. Sin embargo, cuando introducía esa misma palabra poco familiar mientras explicaba un tema poco conocido para los niños, la maestra acompañaba la emisión de la palabra con la realización de gestos.

En el marco de los resultados de las investigaciones reseñadas, el presente estudio se propone analizar el papel que cumplen los gestos de las maestras y de los alumnos durante el proceso de construcción de significados en clases de la escuela primaria. Se asume que los gestos pueden tener un lugar destacado en el aprendizaje de conceptos. Es por ello por lo que el análisis tiene como objetivo atender específicamente a aquellos fragmentos de intercambios en los que se explican, precisan y definen palabras poco frecuentes o desconocidas, así como también cuando los participantes corrigen el uso inapropiado de una palabra, o muestran no comprender una palabra emitida por otro participante e intentan reparar la situación comunicativa.

2. Metodología

El presente estudio adopta una metodología mixta (Bryman, 2006; Cameron, 2011; Creswell, 2011) que combina de modo secuencial una fase cualitativa y una cuantitativa.

En un primer momento, se privilegió la elaboración inductiva de categorías acerca de la articulación entre palabras y gestos en los intercambios conversacionales que configuran las situaciones de enseñanza, siguiendo los procedimientos propuestos por la teoría fundamentada (Glaser y Strauss, 1967; Strauss y Corbin, 1990; Valles, 2003) y su consideración de acuerdo con otros marcos conceptuales (Gumperz, 1982; McNeill, 1985, 1992, 2005; Goodwin, 2000, 2007).

A este fin, se emplearon metodologías audiovisuales que posibilitaron una aproximación ecológica que proporcionó información valiosa sobre las características de los comportamientos verbales y no verbales que conforman las oportunidades de aprendizaje de vocabulario en las situaciones de enseñanza. La metodología usada permitió establecer un orden conceptual en los datos que diera cuenta del objetivo de la investigación, orientado al descubrimiento de la función que los gestos de las maestras y de los niños cumplían en las interacciones.

En un segundo momento, se analizaron cuantitativamente las diferencias en las categorías de gestos empleados por las maestras y los alumnos.

2.1. Participantes

Participaron en este estudio 12 maestras y 368 alumnos de 1.^º (113 niños, la edad promedio era 5:10), 3.^º (108 niños, la edad promedio era 8:2) y 5.^º grado (147 niños, la edad promedio era 10:3) de 4 escuelas públicas primarias de Córdoba, Argentina. Ellos integraban 12 cursos (4 de 1.^º grado, 4 de 3.^º y 4 de 5.^º grado). Los 12 grupos de alumnos procedían de diversos sectores socioeconómicos. Todos los alumnos eran hablantes nativos de español y poseían competencias lingüísticas acordes con su edad.

Los cursos fueron escogidos según criterios de representación típica (Sampieri, Collado, Lucio y Pérez, 1998): se seleccionaron casos que variaran en cuanto al grado de escolaridad y que proporcionaran información abundante sobre los aspectos que estudiar.

Las docentes de todos los grados eran de sexo femenino y su edad oscilaba entre los 26 y los 52 años. En promedio, tenían 15 años de trayectoria en la práctica docente.

2.2. Procedimientos de obtención de información empírica

Se hicieron observaciones videofilmadas de 12 situaciones de enseñanza en los 12 cursos mencionados. Las situaciones de enseñanza registradas en cada curso están conformadas por el desarrollo completo de la unidad temática *Tipos de trabajos* correspondiente al área de ciencias sociales. La información empírica analizada pertenece al *corpus* de nivel primario: Menti, Alejandra (2007-2014). Vocabulario y Aprendizaje. Intercambios Verbales en el Aula. Directora: Rosemberg, Celia. (CONICET). DOI en curso.

No se mencionó a los participantes que se iban a analizar los gestos que realizaran durante el tratamiento del vocabulario. Esto permitió estudiar los gestos espontáneos –no solicitados ni ensayados– por parte de las docentes y de los alumnos.

Las docentes recibieron solo la indicación de llevar a cabo el desarrollo de la unidad temática del mismo modo en que lo harían sin la presencia de la investigadora que filmaba las clases. En consecuencia, las maestras desarrollaron la unidad temática en la cantidad de tiempo que consideraron necesario. Así, la duración total de cada unidad osciló entre 1 y 8 horas cátedra. Todas las situaciones de enseñanza sumaron un total de 36 horas cátedra, integradas por 11.372 turnos de habla (5.013 correspondientes a las docentes y 6.359 a los alumnos).

2.3. Transcripción de los datos

Se efectuó una cuidadosa transcripción literal de las expresiones lingüísticas, como también de la información del contexto situacional y de los comportamientos no verbales que integraron las clases registradas. La descripción de los símbolos utilizados en la transcripción se presenta en el anexo.

2.4. Procedimientos de análisis de la información empírica

En el contexto de las situaciones de enseñanza observadas, se identificaron y delimitaron 72 secuencias interaccionales (Goodwin y Heritage, 1990), en las que docentes y alumnos explicaban, definían o precisaban vocabulario poco frecuente o desconocido, mediante la yuxtaposición de distintos sistemas semióticos (Goodwin, 2000, 2007). En el marco de estas secuencias, se identificaron 226 gestos (171 realizados por las maestras, y 55, por los alumnos).

Se hizo un análisis cualitativo de las secuencias seleccionadas, que combinó el uso del método comparativo constante (Glaser y Strauss, 1967) con el empleo heurístico de conceptos desarrollados por la sociolingüística interaccional (Gumperz, 1982), el análisis conversacional (Goodwin, 2000, 2007) y los estudios de los gestos (McNeill, 1985, 1992, 2005, 2016), considerando su ajuste a los datos.

El diseño metodológico empleado posibilitó analizar, de modo situado, los gestos de las docentes y de los alumnos, codificando la función que cumplen en la construcción colaborativa de los significados de vocabulario poco familiar. Las autoras acordaron en un 91,79 % de los casos en la codificación de las categorías elaboradas. Según lo indicó el coeficiente kappa de Cohen, la concordancia obtenida entre ambas investigadoras fue robusta ($K = 0,880$).

De este modo, el análisis inductivo llevado a cabo permitió elaborar categorías para conceptualizar la función que cumplían los gestos cuando los intercambios conversacionales entre la maestra y los niños se focalizaban en palabras poco familiares, así como también las distintas formas en que los gestos y las intervenciones verbales se yuxtaponían en los intercambios. Dichas categorías se presentan en la tabla 1.

Tabla 1. Categorías elaboradas para dar cuenta del uso de gestos por parte de los docentes y de los alumnos

Categoría	Definición
Ampliar o reforzar información	El interlocutor, mediante sus gestos, proporciona información adicional o la misma que sus expresiones verbales. En este caso, los gestos representan características perceptivas o funcionales del concepto al que se alude.
Cubrir el vacío lingüístico	En su turno de habla, el interlocutor se refiere a determinados conceptos mediante la yuxtaposición del gesto a palabras poco precisas.
Manifestar comprensión compartida	El interlocutor reitera uno o varios gestos producidos por otro hablante en la intervención anterior.
Enfatizar aspectos de sus intervenciones	El interlocutor yuxtapone con intensidad uno o varios gestos a expresiones verbales que quiere destacar.

Fuente: Elaboración propia.

En el apartado de resultados, se presentó primero el análisis cualitativo que dio lugar a la elaboración de estas categorías, de acuerdo con las características metodológi-

cas de los estudios comprensivos, que buscan acercar al lector la realidad concreta que conoció el investigador. El análisis conduce a la elaboración de los datos, respetando su inserción original en un contexto complejo y la presentación de los rasgos más salientes en una descripción que reproduce para el lector su calidad singular (Geertz, 1973; Gibaja, 1988). De este modo, por medio del análisis de ejemplos, de transcripciones de episodios y de citas textuales, se hace una aproximación al ideal de descripción densa, que Geertz (1973) elevó a paradigma del modelo.

Posteriormente, se comparó cuantitativamente la distribución de las categorías elaboradas en el *corpus* de datos. Se establecieron diferencias y similitudes entre las categorías dentro de cada uno de los grupos de interlocutores -docentes y alumnos-, mediante el empleo de la prueba no paramétrica Friedman. Se optó por la aplicación de esta prueba dado que el estadístico Shapiro-Wilk mostró que los valores de las variables incluidas en este análisis no se ajustaban a una distribución normal.

3. Resultados

El análisis de las secuencias interaccionales centradas en el tratamiento del vocabulario puso de manifiesto la función que cumplen los gestos que realizan los alumnos y las maestras cuando se enseña, define, corrige o repara el significado de las palabras poco familiares o desconocidas, como también mostró las distintas formas en que los gestos y las intervenciones verbales se yuxtaponían en los intercambios seleccionados.

3.1. Resultados del análisis cualitativo⁸⁶

En los intercambios entre las maestras y los alumnos durante las situaciones de enseñanza y aprendizaje, los significados se elaboran conjuntamente mediante los aportes de los participantes, en los que se yuxtapone información de distintos campos semióticos. En este sentido, se observó que tanto las maestras como los niños se valen de gestos cuando solicitan un término específico, cuando describen acciones o un objeto en particular, o cuando elaboran definiciones contextuales, como se observa en la siguiente secuencia de intercambio.

Secuencia 1

Esta secuencia conversacional se desarrolló durante los primeros minutos de la clase, en los que una maestra de 1.^º grado comenzó el desarrollo de la unidad temática Tipos de trabajos. Con el objeto de que sus alumnos recuperaran sus conocimientos previos para relacionarlos con el nuevo tema, la maestra les preguntó acerca de los trabajos que realizan sus padres. Como la escuela está ubicada en una zona rural, la mayoría de los niños mencionaron actividades relacionadas con la agricultura. En el siguiente intercambio se observó cómo un niño explicaba las tareas de sembrar y cosechar, focalizándose en las características perceptivas y funcionales de las maquinarias que se emplean para esta actividad productiva.

⁸⁶ Los símbolos empleados en los intercambios conversacionales se encuentran especificados en el Anexo.

1. LAUTARO: las máquinas para semBRÁ(R)LO(.)[¹tienen un cosito que lo TI::ra a la semILLA]

[¹*'Coloca el BD hacia arriba y el BI abajo. Luego, deja caer la mano derecha sobre la mano izquierda, imitando la acción de sembrar'*]



Figura 1.1

Figura 1.2

Figura 1.3

Figura 1 "Tienen un cosito que lo tira a la semilla"

después pasa una [²COSITA] () [³que lo hace así [⁴MONTAÑITA]

[²*Extiende hacia adelante ambos brazos; une los dedos de ambas manos y forma un triángulo, representando el cierre del surco]*

[³*Coloca ambos antebrazos en el banco, con las manos extendidas. Los mueve hacia adelante hasta formar un triángulo, representando el movimiento de la sembradora cuando va tapando el surco y, además, la forma con la que queda el surco tapado]*



Figura 2



Figura 3

Figura 2 "Cosita" Figura 3 "Que lo hace así montañita"

la /COSECHADORA () como yo vi una vez (www) () [⁴y /saca]

[⁴Mueve ambas manos —MD con más énfasis— de abajo hacia arriba, imitando la acción de sacar]

(...)

la semillita [⁵se va un lado grande]

[⁵Con ambos brazos dibuja un amplio círculo arriba de su cuerpo, representando la tolva de la cosechadora]



Figura 4.1



Figura 4.2

Figura 4 “Se va un lado grande”

y después (.)[⁶saca un /COSO] (.)[⁷viene la /TOLVA] (.)

[⁶Se toca el pecho con la MD y luego extiende el antebrazo hacia adelante, imitando el movimiento que realiza el tubo de la cosechadora]

[⁷Dibuja con ambas manos un círculo, representando la tolva]



Figura 5

Figura 6

Figura 5 “Saca un coso” **Figura 6** “Viene la tolva”

(www) para que [⁸/SEPAN]

[⁸*Junta los dedos de cada mano y realiza un movimiento corto y preciso con ambas manos de arriba hacia abajo*]



Figura 7 “Sepan”

quién es la tolva (.) [⁹*tiene un*] [¹⁰*(www) (.) unas lucecitas que hacen yin yin yin*]

[⁹*Coloca la MDencima de su cabeza y la mueve de arriba hacia abajo*]

[¹⁰*Coloca ambas manos enfrentadas a la altura de su cabeza y las abre y cierra al compás de la onomatopeya*]



Figura 8 “Tiene un”



Figura 9.1



Figura 9.2

Figura 9 “Unas lucecitas que hacen yin yinyin”

(...)

2. MAESTRA: muy bien (.) con la /TOLva van /cargando (.) a::: (.) todas la::s /semillas que van cosechando.

Tal como fue señalado en estudios anteriores (Crowder y Newman, 1993; Newman y Crowder, 1991; Lemke, 1999; Roth, 2000, 2002), en la secuencia de interacción se observó que los alumnos de 1.^º grado se valían no solo de expresiones lingüísticas, sino también de gestos cuando manifestaban el conocimiento que poseen acerca de un tópico. A lo largo de su explicación, el alumno recurría a gestos, en su mayoría icónicos, para transmitir información cuando no disponía o no podía acceder fácilmente a las palabras que necesitaba para explicar los procesos de siembra y cosecha. El alumno producía gestos icónicos, ya sea en co-ocurrencia con palabras poco precisas que refieren a partes específicas de la sembradora y de la cosechadora -gesto 1*cosito*, gesto 2*cosita*, gesto 5 *lado*, gesto 6*coso*-; para reemplazar palabras -gesto 9(*Coloca la MD encima de su cabeza y la mueve de arriba hacia abajo representando así el tubo de la tolva*); gesto 10(*Coloca ambas manos enfrentadas a la altura de su cabeza y las abre y cierra al compás de la onomatopeya. De esta manera, representan el sonido que emite la tolva cuando está llena*)-; para ampliar o reforzar información proporcionada a través de sus expresiones lingüísticas -gesto 1*tira*, gesto 4*saca*, gesto 5 *grande*, gesto 7*tolva*-; o bien cuando explicaba un concepto mediante comparaciones por medio de un referente más conocido, a la manera de una metáfora -*que lo hace así montañita*, gesto 3(*Coloca ambos antebrazos en el banco, con las manos extendidas. Los mueve hacia adelante hasta formar un triángulo, representando el movimiento de la sembradora cuando va tapando el surco y, además, la forma en que queda el surco tapado*)-.

De este modo, se pone de manifiesto cómo el empleo de gestos icónicos por parte del niño puede funcionar como un modo de expresión alternativo o complementario, que tiende a cubrir el vacío lingüístico que se produce al no tener disponible en su léxico mental la palabra específica que necesita para manifestar su conocimiento acerca de un tema en particular (Newman y Crowder, 1991; McNeill, 1992; Crowder y Newman, 1993; Lemke, 1999; Roth, 2000).

Secuencia 2

En el siguiente fragmento, se presenta una secuencia en la que una maestra de 3.^º grado elaboró conjuntamente con sus alumnos una definición contextual de la palabra “silo”. Para ello, se focalizaron principalmente en las características funcionales y perceptivas del referente:

1. MAESTRA: ¿A ver qué características tiene la zona rural (.) que no tiene la zona urbana? =a ver=
2. ((Varios niños levantan la mano))
3. MAESTRA: a ver ((Señala a la niña que tiene el turno para hablar))
4. GUADALUPE: calles de tierra.
5. MAESTRA: calles la mayoría de tierra.
6. AGUSTINA: animales.
7. MAESTRA: muchos animales.
8. MAXI: silos
9. MAESTRA: silos (.) / ¿QUÉ ES SILOS?
10. JOSÉ: =[¹¹/es como una cosa]=

[¹¹Extiende rápidamente ambas manos hacia arriba]



Figura 10

Figura 10 “Es una cosa”

MAXI: =seño=

11. MAESTRA: ¿a ver ?
12. NIÑOS: (www) ((Hablan todos juntos))
13. MAESTRA: / PAREN (.) PAREN
14. MAXI: son::: (.) [¹²/no sé cómo explicarlo] (.)

[¹²Se tapa la cara con las palmas de ambas manos]



Figura 11

Figura 11 “No sé cómo explicarlo”

16. JOSÉ: son

17. LUCAS: seño (.) =¡Yo YoYo !=
18. MAESTRA: =PAREN (.) PAREN= ya lo va a explicar\
19. JOSÉ: YO YOYOYO
20. MAXI: [¹³/son grandes]

[¹³Mueve la MD de arriba hacia abajo mientras dibuja una S en el aire]



Figura 12.1

Figura 12.2

Figura 12.3

Figura 12 “Son grandes”

21. LUCAS: y arriba tienen [¹⁴como una cosa]



Figura 13

Figura 13 “Como una cosa”

[¹⁴Forma un triángulo con ambas manos al unir los dedos índice, representando el techo del silo]

22. MAESTRA: =¿pero qué son?=.
 23. JOSÉ: =((levanta la mano))=.
 24. MAESTRA: José ((cede el turno a José))
 25. JOSÉ: [¹⁵eh (.) es (.) es como un pote] [¹⁶pero regrande /así]

[¹⁵Con ambas manos dibuja el contorno de un recipiente muy grande]



Figura 14

Figura 14 “Eh, es es como un pote”

[¹⁶Extiende ambos antebrazos hacia los costados, representando el tamaño del silo]



Figura 15

Figura 15 “Pero regrande, así”

26. MAESTRA: como [¹⁷unos frascos ENO::R::MES] (.) un RECIPIENTE grande

[¹⁷Levanta ambos brazos sobre su cabeza]



Figura 16

Figura 16 “Unos frascos ENO::R::MES”

Este fragmento de conversación se situó al comienzo de la clase. A modo de introducción de la unidad temática Tipos de trabajos, la maestra pidió a sus alumnos que mencionaran características que diferencian a la zona urbana de la zona rural. En el turno 8, uno de los alumnos mencionó la palabra "silo". Cabe suponer que la maestra asume un conocimiento parcial por parte de los niños de lo que ello es. Esto daría cuenta del hecho de que en el turno siguiente iniciara un movimiento conversacional en el que, mediante una pregunta, solicitaba la definición del término.

Las respuestas de los alumnos ponen de manifiesto que la palabra foco "silo" era un vocablo poco familiar para ellos, puesto que, si bien parecían saber a qué refiere la palabra, no podían definirla. En el turno 10, por ejemplo, un niño intentó definirla por medio de una comparación con un término poco preciso -*es como una cosa*-. Dado que el término al que recurrió resulta muy vago y general -*cosa*-, es probable que el niño haya percibido que su definición no estaba completa; de ahí que acompañara su respuesta con un gesto icónico -gesto 11(*Extiende rápidamente ambas manos hacia arriba*)- que le permitió representar la forma del silo.

Cuando la maestra dio la palabra a otro alumno, en el turno 15, el niño interrumpió su emisión, puesto que en este caso tampoco parecía disponer de las palabras precisas para dar cuenta de qué es un silo. Hizo una breve pausa y luego empleó un gesto metafórico -gesto 12(*se tapa la cara con un mano*)- al mismo tiempo que señaló su dificultad para definir el concepto -*no sé cómo explicarlo*-.

En el turno 20, este mismo alumno realizó otro intento. En este caso, yuxtapuso otro gesto icónico -gesto 13(*Mueve la MD de arriba hacia abajo mientras dibuja una S en el aire*)- a las palabras que pronunciaba. Mientras que su expresión verbal se focalizaba en las características perceptuales del referente -*son grandes*-, el gesto icónico aludía a sus características funcionales. En efecto, el dibujo de la S en el aire representaba la acción de depositar cereal en el silo. Por su parte, otro alumno interrumpió su intervención para completar la respuesta de su compañero, focalizándose en otra característica perceptiva del silo, el techo, mediante la yuxtaposición de campos semióticos -verbal y gestual-. Al igual que sus compañeros, solo disponía de una palabra vaga e imprecisa para definirlo -*tiene como una cosa*-, pero reforzó su respuesta valiéndose de un gesto icónico que representaba la forma del techo de un silo -gesto 14(*Forma un triángulo con ambas manos, al unir los dedos índice, representando el techo del silo*)-.

La maestra, aparentemente no conforme con las respuestas, insistió, en el turno 22, con la misma pregunta que formuló en el turno 8 -*¿pero qué son?*- mediante la que solicitaba una definición que diera cuenta del concepto.

En el turno 25, otro niño respondió estableciendo una comparación con otro referente más conocido -*es como un pote, pero regrande, así*- que posee características perceptuales y funcionales similares al silo, pero a menor escala. Su respuesta verbal fue complementada con dos gestos icónicos: uno que representaba la forma del silo -gesto 15(*Con ambas manos dibuja el contorno de un recipiente muy grande*)-, y el otro, su tamaño -gesto 16(*Extiende ambos antebrazos hacia los costados*),

representando el tamaño del silo)-. Esta respuesta se aproximaba, en mayor medida, a la respuesta esperada por la maestra.

Consecuentemente, en el turno siguiente esta retomó la respuesta del niño -*es como un pote, pero grande, así*- y la reformuló valiéndose de dos estructuras sintácticas isomórficas: en la primera, la maestra reemplazó las palabras léxicas pronunciadas por el niño -*pote, grande*- por palabras algo menos familiares -*frascos, ENO::R::MES*-. Esta reformulación se encontró acompañada por un gesto icónico -gesto 17(*Levanta ambos brazos sobre su cabeza*)- que representaba el tamaño de los silos, cualidad aludida desde el primer momento por los alumnos. El gesto, a su vez, y el alargamiento vocálico del término se reforzaban mutuamente entre sí, para enfatizar la característica expresada mediante esa palabra -*ENO::R::MES*-. Mientras que, en la segunda estructura, que es isomórfica con la primera, la maestra ofrecía, con mucho énfasis, un término aún más específico y menos familiar -*RECIPIENTE grande*-. En esta instancia, ella trasladó, por un proceso de reconceptualización progresiva -*pote, frasco, recipiente*-, el foco al concepto que resultó problemático a lo largo de la secuencia presentada.

En este intercambio, la reparación, que implicó una reconceptualización progresiva, dio cuenta de la organización secuencial de la interacción (Schegloff, Jefferson y Sacks, 1977; Macbeth, 2004). En efecto, a partir de las respuestas de sus alumnos, la maestra reconoció que podía producirse un problema en la comunicación. Las dificultades que mostraron al definir la palabra foco, "silo", podían traer aparejado que otros niños en la situación de enseñanza no comprendieran

fehacientemente el significado del término. En consecuencia, ella generó una serie de movimientos destinados a reparar la situación comunicativa.

Así, se muestra que los alumnos apelan a gestos icónicos para representar de manera visual lo que no pueden hacer verbalmente, ya sea porque no disponen de las palabras precisas o debido al conocimiento parcial que poseen de la palabra (Newman y Crowder, 1991; McNeill, 1992; Crowder y Newman, 1993; Wells, 1998; Lemke, 1999; Roth, 2000).

Resulta interesante señalar que la maestra, durante el proceso de reconceptualización progresiva que llevó adelante, repitió, con algunas variantes, el gesto icónico realizado por el alumno que formuló la respuesta más aproximada a la esperada. Esta repetición gestual podría haber cumplido diversas funciones: por un lado, le permitió evaluar positivamente la respuesta del alumno y reducir la distancia interaccional (Mehrabian, 1967), retomando en su emisión parte de la producida por el alumno; y, por otro, tal como fue señalado por Roth (2001), le proporcionó un apoyo o recurso adicional que reforzó lo transmitido.

Secuencia 3

En el siguiente intercambio conversacional que tuvo lugar durante una clase en 5.^º grado, se muestra cómo la información lingüística y gestual que proporcionó la docente posibilitó que una alumna infiriera el nombre propio solicitado por la docente. Asimismo, se observa claramente cómo el intercambio de significados entre los interlocutores se produjo no solo a nivel lingüístico, sino también a nivel gestual:

1. MAESTRA: pero hay [¹⁸una provincia que está pegadi::ta] [¹⁹a /Córdoba]

[¹⁸Mueve la MI de arriba hacia abajo. Con los dedos forma una "U", mientras realiza el movimiento, representado así la forma rectangular de la provincia de San Luis]

[¹⁹Mueve la MI muy sutilmente hacia adelante y hacia atrás]



Figura 17

Figura 17 “Una provincia que está pegadi::ta”



Figura 18

Figura 18 “A Córdoba”

2. ALUMNOS: ((varios alumnos levantan la mano para contestar))

3. MATÍAS: /_cSantiago del Estero?

4. LUCAS: San Luis

5. MAESTRA: [²⁰((silencio)) ¿por] [²¹qué /San Luis? (.) /a ver]

[²⁰Señala a la alumna con el dedo índice de la MD]

[²¹Se toma ambas manos y las lleva hacia su mentón. Mira a la alumna mientras sonríe]



Figura 19

Figura 19 “((Silencio)) ¿Por...?”



Figura 20

Figura 20 “¿..qué San Luis?”

6.LUCÍA: e::h[²²usted hizo /así]

[²²Imita el gesto 18 que hizo la maestra, pero con algunos cambios. Coloca ambas manos a la altura de su rostro con ambos dedos índice, mirando hacia la maestra, y mueve las manos de arriba hacia abajo, representando la forma rectangular de la provincia de San Luis]



Figura 21.1

Figura 21 “Usted hizo así”



Figura 21.2

7. MAESTRA: que [²³está claro]

[²³Repite gesto 18, pero con un movimiento más breve]



Figura 22.1



Figura 22.2

Figura 22 “Está claro”

[²⁴Córdoba]

[²⁴Sitúa la MD a la altura de su rostro con todos los dedos de su mano mirando hacia arriba]



Figura 23
Figura 23 Córdoba”

[²⁵y la tenemos al lado]

[²⁵Mueve la MI de arriba hacia abajo. Con los dedos forma una “U”, mientras realiza el movimiento]



Figura 24

Figura 24 “Y la tenemos al lado”

[²⁶como que está en la /espalda] de Córdoba

[²⁶Se toca la espalda con la MD]



Figura 25

Figura 25 “Como que está en la espalda”

En el intercambio, la alumna se apoyó en la información lingüística -*una provincia que está pegada:ta a Córdoba*-,⁸⁷ pero sobre todo gestual -gesto 18 (*Mueve la MI de arriba hacia abajo. Con los dedos forma una U, mientras realiza el movimiento, representando así la forma rectangular de la provincia de San Luis*)-, que le proporcionó la maestra en el turno 1 para interpretar a qué provincia está haciendo referencia -San Luis-. Esto fue puesto de manifiesto, explícitamente, por la niña en el turno 6, cuando expresó que el gesto metafórico realizado por la docente -gesto 18- ofreció pistas que orientaron su inferencia -*Usted hizo /así/*-.

Cabe señalar que, en este caso, el intercambio completo de significados se produjo no solo a nivel verbal, sino también gestual. En efecto, la alumna, mientras decía *así*, imitaba el gesto de la maestra -gesto 18- con algunas variaciones -gesto 22 (*coloca ambas manos a la altura de su rostro con ambos dedos índice, mirando hacia la maestra, y mueve las manos de arriba hacia abajo, representando la forma rectangular de la provincia de San Luis*)-. A modo de respuesta, la maestra destinó la siguiente intervención -turno 7- a explicar y ampliar la información gestual proporcionada en el turno 1. Para ello, se apoyó en distintos gestos metafóricos mediante los que representó la forma de la provincia de San Luis -gesto 23-, la ubicación de la provincia de Córdoba con respecto a San Luis -gesto 24-, la ubicación de ambas provincias -gesto 25-, y la zona de Córdoba que limita con San Luis -gesto 26-, con la intención de delimitar la ubicación geográfica

⁸⁷ Provincia de la República Argentina ubicada en el centro del país. Limita con siete provincias: al norte con Catamarca y Santiago del Estero; al este con Santa Fe; al sureste con Buenos Aires; al sur con La Pampa y al oeste con San Luis y La Rioja (Gobierno de la Provincia de Córdoba, s. f.).

de ambas provincias a las que aludía en el intercambio. Como lo señalaron Singer y Goldin-Meadow (2005), en esta secuencia de interacción se observa cómo los gestos del docente que transmiten información distinta a la expresada en su discurso ofrecen a los alumnos un camino alternativo para resolver problemas que se plantean en la clase.

3.2. Resultados del análisis cuantitativo

Para dar cuenta cuantitativamente de la función que cumplen los gestos de los interlocutores -docentes y alumnos- en las situaciones en las que se enseñaba, definía, corregía o reparaba el significado de las palabras poco familiares o desconocidas, se llevó a cabo un análisis intragrupal. Este análisis mostró cómo se distribuyen los gestos de las maestras, por un lado, y de los alumnos, por otro, según la función que cumplen en aquellas situaciones en las que se enseña vocabulario. Los resultados se presentan en las tablas 2 y 3.

Tabla 2. Gestos de las maestras

Función	Suma (Rangos)	Media (Rangos)	N	T ²	p
Comprensión compartida	23,50	1,96	12 A	8,77	0,0002
Cubrir el vacío lingüístico	26,50	2,21	12 A B		
Enfatizar	28,50	2,38	12 A B C		
Ampliar o reforzar información	41,50	3,46	12 D		
<i>Medias con una letra común no son significativamente diferentes ($p > 0,050$)</i>					
N: Cantidad de cursos observados; T ² : Estadístico calculado del análisis de la varianza por rangos de Friedman; p: Medida de significación estadística					

Fuente: Elaboración propia.

Como se observa en la tabla 2, durante los intercambios conversacionales en los que se enseñan palabras poco familiares o desconocidas a los alumnos, las maestras tienden a yuxtaponer a sus explicaciones lingüísticas, en mayor medida, gestos que tienen la función de ampliar o reforzar la información que transmiten verbalmente. En este caso, la prueba de Friedman evidenció diferencias estadísticamente significativas entre las categorías ($T^2(N = 12) = 8,77, p = 0,0002$).

Tabla 3. Gestos de los alumnos

Función	Suma (Rangos)	Media (Rangos)	N	T ²	p
Enfatizar	24,50	2,04	12 A	4,78	0,0071
Comprensión compartida	26,50	2,21	12 A B		
Cubrir el vacío lingüístico	30,00	2,50	12 A B C		
Ampliar o reforzar información	39,00	3,25	12 D		

Medias con una letra común no son significativamente diferentes ($p > 0,050$)

Fuente: Elaboración propia.

Por su parte, la tabla 3 muestra la distribución de los gestos realizados por los alumnos según la función que cumplen en aquellas situaciones en que se enseña vocabulario. En esta tabla se observa que, en el caso de los alumnos, priman aquellos gestos destinados a ampliar o reforzar la información que proporcionan verbalmente por sobre los demás tipos de gestos y, en segundo lugar, los gestos destinados a cubrir el vacío lingüístico. Al igual que en el caso anterior, la aplicación de la prueba de Friedman puso de manifiesto diferencias significativas ($T^2(N=12) = 4,78, p = 0,0071$).

4. Conclusiones

Los resultados del análisis cualitativo de los intercambios entre docentes y alumnos pusieron de manifiesto el importante rol que cumplen los gestos en ellos y el modo en que

se superponen a las expresiones verbales, cuando docentes y niños introducen, explican, precisan, definen o corrigen palabras poco frecuentes o desconocidas durante clases de ciencias sociales. En efecto, los resultados del presente trabajo mostraron que todos los interlocutores de las secuencias de interacción registradas -maestras y niños- emiten gestos cuando solicitan un término específico, describen acciones o un objeto en particular, y cuando elaboran definiciones contextuales de manera conjunta.

Aun cuando maestras y alumnos realizan los mismos gestos durante el tratamiento de palabras poco familiares o desconocidas, la función que desempeñan los gestos de unos y otros varía dentro de la secuencia de interacción. Como se mostró en el análisis cuantitativo, se observan diferencias significativas en las funciones para las cuales emplean los gestos tanto las maestras como los alumnos. El análisis cuantitativo mostró que los interlocutores -docentes y alumnos- privilegian el empleo de gestos destinados a ampliar o reforzar la información que proporcionan verbalmente por sobre los demás tipos de gestos.

Se observó, además, que los alumnos tienden a emplear gran cantidad de gestos icónicos cuando no disponen o no pueden acceder fácilmente a las palabras que necesitan para explicar un tema en particular (Newman y Crowder, 1991; McNeill, 1992; Crowder y Newman, 1993; Lemke, 1999; Roth, 2000). En este caso, los gestos de los alumnos les permiten adoptar una modalidad de expresión alternativa que sirve para cubrir el *vacío lingüístico* (McNeill, 1992; Crowder Newman, 1993; Wells, 1998; Lemke, 1999; Krauss, Chen y Gottesman, 2000; Roth, 2000, 2001).

Por su parte, los docentes tienden a repetir los gestos icónicos efectuados por los alumnos. Esta repetición gestual cumple, en cambio, distintas funciones: valorar como

positiva la respuesta del alumno y reducir la distancia en la interacción (Mehrabian, 1967), al recuperar, en su turno de habla, parte de lo expresado por aquel y servir de apoyo o recurso adicional que refuerza lo transmitido.

Cuando la maestra reitera en su intervención un gesto realizado por el niño, ella implícitamente muestra que ha comprendido lo que el niño quería trasmitir, y, en el marco de ese anclaje de comprensión mutua, la maestra puede continuar proporcionando información adicional.

Los docentes también se apoyan en la yuxtaposición de diferentes sistemas semióticos (Goodwin, 2000, 2007), verbal y no verbal, que expresan distinta información para obtener contenidos específicos por parte de sus alumnos. Al igual que lo señalado en el estudio de Singer y Goldin-Meadow (2005), en este caso se observa cómo los gestos del docente que transmiten información distinta a la expresada en su discurso ofrecen a los alumnos una vía alternativa para resolver problemas que se plantean en la clase y, en consecuencia, redundan en beneficio para el aprendizaje de los alumnos.

Los resultados de este estudio tienen importantes implicancias pedagógicas si se consideran en el marco de investigaciones previas que destacaron la relevancia de las interacciones entre el adulto y el niño para el aprendizaje de palabras (Beals y Tabors, 1995; Nelson, 1996, 2007; Tomasello, 1998, 2003; Weizman y Snow, 2001). Como señalaron estos trabajos, el aprendizaje de vocabulario depende no solo de la cantidad, la diversidad y la complejidad de las palabras que el adulto presenta al niño, sino también del andamiaje social y pragmático que padres y maestros proporcionan durante la situación comunicativa. En este sentido, el presente estudio pone de manifiesto la necesidad de que el docente atienda no solo a la información verbal que le proporciona al alumno cuando

solicita, introduce o define un término desconocido o poco familiar, sino también al empleo de gestos que le permiten al niño inferir un término en particular o comprender el significado de las palabras que desconoce o conoce de modo parcial.

Los gestos constituyen un sistema semiótico de gran relevancia en las situaciones de enseñanza de ciencias, no solo naturales, sino también sociales, puesto que en ellos se apoyan los docentes para desarrollar el discurso científico (Roth, 2000, 2002; Roth y Welzel, 2001; Roth y Lawless, 2002), especialmente cuando explican conceptos abstractos (Roth y Welzel, 2001; Roth y Lawless, 2002; McNeill, 2005). Por su parte, los gestos de los alumnos ayudan a que estos expresen sus conocimientos, en particular cuando no disponen o no pueden acceder a la palabra precisa (Wells, 1998; Roth, 2000, 2002; Roth y Welzel, 2001; Roth y Lawless, 2002, entre otros).

Referencias bibliográficas

- Babad, E., Avni-Babad, D. y Rosenthal, R. (2004). "Prediction of students' evaluations from brief instances of professors' nonverbal behavior in defined instructional situations". *Social Psychology of Education*, 7(1), pp. 3-33.
- Barowy, W. y Smith, J. (2008). "Ecology and development in classroom communication". *Linguistics and Education*, 19, pp. 149-165.
- Beals, D. y Tabors, P. (1995). "Arboretum, bureaucratic and carbohydrates: Preschoolers' exposure to rare vocabulary at home". *First Language*, 15, pp. 57-76.

- Bryman, A. (2006). "Integrating quantitative and qualitative research: How is it done?". *Qualitative Research*, 6(1), pp. 97-113.
- Cameron, R. (2011). "Mixed methods research: The five Ps framework". *The Electronic Journal of Business Research Methods*, 9(2), pp. 96-108.
- Cook, S. W., Mitchell, Z. y Goldin-Meadow, S. (2008). "Gesturing makes learning last". *Cognition*, 106(2), pp. 1047-1058.
- Creswell, J. (2011). "Controversies in mixed methods research". En N. K. Denziny y S. Lincoln (Eds.), *The Sage Handbook of Qualitative Research*, (4.aed., pp. 269-283).
- Crowder, E. M. y Newman, D. (1993). "Telling what they know: The role of gesture and language in children's science explanations". *Pragmatics y Cognition*, 1(2), pp. 341-376.
- Edwards, L. (2009). "Gestures and conceptual integration in mathematical talk". *Educational Studies in Mathematics*, 70(2), pp. 127-141.
- Furuyama, N. (2000). "Gestural interaction between the instructor and the learner in origami instruction. En D. McNeill (Ed.), *Language and Gesture* (pp. 99-117). Cambridge: Cambridge University Press.
- Gaythwaite, E. S. (2005). "Didn't you see what I meant?". *Curriculum and Teaching Dialogue*, 7(1-2), pp. 97-108.
- Geertz, C. (1973). *The Interpretation of Cultures: Selected Essays*. Nueva York: Basic Books.
- Gibaja, R. (1988). "Acerca del debate metodológico en la investigación educacional. La educación". *Revista Interamericana de Desarrollo Educativo (OEA)*, 103, pp. 81-94.
- Glaser, B. G., y Strauss, A. L. (1967). *The Discovery of Grounded Theory*. Chicago: Aldine Publishing Company.

- Gobierno de la Provincia de Córdoba (s. f.). Provincia. Ubicación. Recuperado de <https://bit.ly/2uz9kRg>.
- Goldin-Meadow, S. (2003). *Hearing Gesture: How Our Hands Help Us Think*. Cambridge, MA: Harvard University Press.
- Goldin-Meadow, S. (2006). "Talking and thinking with our hands". *Psychological Science*, 15(1), pp. 34-39.
- Goldin-Meadow, S. (2014). "How gesture works to change our minds". *Trends in Neuroscience and Education*, 3(1), pp. 4-6.
- Goldin-Meadow, S. y Singer, M. A. (2003). "From children's hands to adults' ears: Gesture's role in teaching and learning". *Developmental Psychology*, 39(3), pp. 509-520.
- Goldin-Meadow, S., Kim, S. y Singer, M. (1999). "What the teacher's hands tell the student's mind about math". *Journal of Educational Psychology*, 91(4), pp. 720-730.
- Goodwin, C. (2000). "Action and embodiment within situated human interaction". *Journal of Pragmatics*, 32, pp. 1489-1522.
- Goodwin, C. (2007). "Participation, stance and affect in the organization of activities". *Discourse y Society*, 18(1), pp. 53-73.
- Goodwin, C.y Heritage, J. (1990). "Conversation analysis". *Annual Review of Anthropology*, 19, pp. 238-307.
- Gumperz, J. (1982). *Language and Social Identity*. Cambridge: Cambridge University Press.
- Kelly, S. D., Özyürek, A. y Maris, E. (2010). "Two sides of the same coin: Speech and gesture mutually interact to enhance comprehension". *Psychological Science*, 21(2), pp. 260-267.

- Kendon, A. (1980). "Gesticulation and speech: Two aspects for the process of the utterance". En M. R. Key (Ed.), *The Relation Between Verbal and Non-Verbal Communication* (pp. 207-227). La Haya: Mouton and Co.
- Kendon, A. (2000). "Language and gesture: Unity or duality?". En D. McNeill (Ed.), *Language and Gesture* (pp. 47-63). Cambridge: Cambridge University Press.
- Krauss, R. K., Chen, Y. y Gottesman, R. F. (2000). "Lexical gestures and lexical access: A process model". En D. McNeill (Ed.), *Language and Gesture* (pp. 261-283). Cambridge: Cambridge University Press.
- Lemke, J. (1998). "Multimedia literacy demands of the scientific curriculum". *Linguistics and Education*, 10(3), pp. 247-271.
- Lemke, J. L. (1999). "Typological and topological meaning in diagnostic discourse". *Discourse Processes*, 27(2), pp. 173-185.
- Levinson, S. (1997). "From outer to inner space: Linguistic categories and non linguistic thinking. En J. Nuyts y E. Pederson (Eds.), *Language and Conceptualization* (pp. 13-45). Cambridge: Cambridge University Press.
- Macbeth, D. (2004). "The relevance of repair for classroom correction". *Language in Society*, 33(5), pp. 703-736.
- McNeill, D. (1985). "So you think gestures are nonverbal?". *Psychological Review*, 92, pp. 350-371.
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. Chicago: University of Chicago Press.
- McNeill, D. (2005). *Gesture and Thought*. Chicago: University of Chicago Press.
- McNeill, D. (2016). *Why We Gesture. The Surprising Role of Hand Movements in Communication*. Cambridge: Cambridge University Press.

- Mehrabian, A. (1967). "Orientation behaviors and nonverbal attitude communication". *Journal of Communication*, 17(4), pp. 324-332.
- Menti, A. y Alam, F. (2014). "Los gestos y la enseñanza de palabras poco familiares: ¿cuándo emplean las maestras información gestual?". *Bellaterra Journal of Teaching y Learning Language y Literature*, 7(1), pp. 17-32.
- Migdalek, M. J. y Rosemberg, C. R. (2012). "El uso de los gestos en el discurso docente durante la planificación del juego en el jardín de infantes". *Bellaterra Journal of Teaching y Learning Language y Literature*, 5(3), pp. 25-43.
- Mondada, L. (2001). "Por una lingüística interaccional". *Discurso y Sociedad*, 3(3), pp. 61-89.
- Nelson, K. (1996). *Language in Cognitive Development. The Emergence of the Mediated Mind*. Cambridge: Cambridge University Press.
- Nelson, K. (2007). *Young Minds in Social Worlds. Experience, Meaning and Memory*. Cambridge: Harvard University Press.
- Newman, D. y Crowder, E. (1991). "Teacher support for sense-making in classroom conversation". *Annual Meeting of the American Educational Research Association*. Chicago.
- Orellana, M., Martínez, D., Lee, C. y Montano, E. (2012). "Language as a tool in diverse forms of learning". *Linguistics and Education*, 23, pp. 373-387.
- Roth, W. M. (2000). "From gesture to scientific language". *Journal of Pragmatics*, 32(11), pp. 1683-1714.
- Roth, W. M. (2001). "Gestures: Their role in teaching and learning". *Review of Educational Research*, 71, pp. 365-392.

- Roth, W. M. (2002). "Reading graphs: Contributions to an integrative concept of literacy". *Journal of Curriculum Studies*, 34, pp. 1-24.
- Roth, W. M. y Lawless, D. (2002). "Scientific investigations, metaphorical gestures and the emergence of abstract scientific concepts". *Learning and Instruction*, 12, pp. 285-304.
- Roth, W. M. y Welzel, M. (2001). "From activity to gestures and scientific language". *Journal of Research in Science Teaching*, 38(1), pp. 103-136.
- Sampieri, R.H., Collado, C.F., Lucio, P.B.y Pérez, M.L.C. (1998). *Metodología de la investigación*. México DF, México: McGraw Hill.
- Schegloff, E. A., Jefferson, G. y Sacks, H. E. (1977). "The preference for self-correction in the organization of repair in conversation". *Language*, 53(2), pp. 361-382.
- Singer, M. A.y Goldin-Meadow, S. (2005). "Children learn when their teacher's gestures and speech differ". *Psychological Science*, 16, pp. 85-89.
- Strauss, A., y Corbin, J. (1990). *Basics of Qualitative Research Techniques and Procedures for Developing Grounded Theory*. Londres: Sage Publications.
- Tomasello, M. (1998). "Introduction: A cognitive-functional perspective on language structure". En M. Tomasello (Ed.), *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure* (pp. vii-xxiii). Mahwah, NJ: Lawrence Erlbaum.
- Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition*. Massachusetts: Harvard University Press.
- Valles, M. (2003). *Técnicas cualitativas de investigación social*. Madrid: Síntesis.

- Weizman, Z. O. y Snow, C. E. (2001). "Lexical input as related to children's vocabulary acquisition: Effects of sophisticated exposure and support for meaning". *Developmental Psychology, 37*(2), pp. 265-279.
- Wells, G. (1998). "Modes of meaning in a science activity". *Linguistics and Education, 10*(3), pp. 307-334.

Anexo

Pautas de transcripción

::	Alargamiento vocálico
/	Tono ascendente
	Tono descendente
(())	Fenómenos no léxicos
MAYUSC	Énfasis
= =	Superposición
(.)	Pausa breve, menor a 0,1 segundos
<0,5>	Pausa extensa (en segundos)
(www)	Inaudible
(dudosos)	Fragmento dudoso
? ?	Entonación interrogativa
-	Interrupción
Γ	Tono agudo
└	Tono grave
[]	Inicio y final del gesto respectivamente
<i>Palabras en itálica</i>	Descripción del gesto
BD	Brazo derecho
BI	Brazo izquierdo
MD	Mano derecha
MI	Mano izquierda

Sobre los autores

Irene Audisio (irene.audisio@gmail.com)

Es doctora en Filosofía por la Universidad Nacional de Córdoba (2011). Actualmente, se desempeña como profesora titular en la cátedra Problemáticas filosóficas de la Facultad de Lenguas (UNC) y como profesora asistente en la Facultad de Filosofía y Humanidades (UNC). Integra el proyecto de investigación “Variedades de la intencionalidad: contenido, expresión y atribución” (Secyt, UNC), dirigido por la Dra. Carolina Scotto y la Dra. Laura Danón del Centro de Investigaciones de la Facultad de Filosofía y Humanidades. Sus temas de interés giran alrededor de los enfoques corporizados de la mente y sus alcances para el estudio del lenguaje. Últimamente, se ha concentrado en los abordajes teóricos sobre gestos.

Xilenia María Carreras (carrerasxilenia@gmail.com)

Lic. en Psicología. Doctoranda del doctorado en Psicología de la Universidad Nacional de Córdoba. Secretaría de Ciencia y Técnica (SeCyT). Trabaja en proyecto doctoral “Reconocimiento del propio cuerpo, agentividad y destrezas narrativas: un estudio exploratorio sobre aquellas competencias asociadas con el desarrollo de la vivencia de mismidad en niños de desarrollo típico y en niños con diagnóstico de TGD”. Miembro del Equipo Interdisciplinario de Atención Integral de los Trastornos del Desarrollo. Se desempeña como terapeuta en el tratamiento de

pacientes del espectro autista en CETES (Centro de Estudio y Tratamiento de la Epilepsia y el Sueño), Fundación Fetení.

Laura Danón (lidanon@gmail.com)

Es licenciada en Psicología y doctora en Filosofía por la UNC. Es profesora de la Facultad de Filosofía y Humanidades y en la Facultad de Psicología de la UNC. Se especializa en filosofía de la mente y filosofía de la psicología. Ha publicado artículos en temas de su especialidad en distintas revistas nacionales e internacionales, tales como *Philosophia*, *Teorema*, *Critica*, *Análisis Filosófico*, *Studia Philosophica Estónica*, *Areté*, etc. Ha editado el libro *Conceptos, lenguaje y cognición* junto con la Dra. Carolina Scotto y la Dra. Mariela Aguilera (UNC, 2015). Codirige, junto con la Dra. Carolina Scotto, el proyecto “Variedades de la intencionalidad: contenido, expresión y atribución”, subsidiado por SeCyT de la UNC.

Daniel Kalpokas (dkalpokas@gmail.com)

Daniel Kalpokas es doctor en Filosofía por la Universidad de Buenos Aires (2004). Actualmente, se desempeña como investigador independiente del CONICET y como profesor adjunto regular en la Facultad de Filosofía y Humanidades de la Universidad Nacional de Córdoba. Es autor de *Richard Rorty y la superación pragmatista de la epistemología* (2005) y de varios artículos publicados en revistas tales como *Diánoia*, *Análisis Filosófico*, *Contemporary Pragmatism*, *Nordic Wittgenstein Review*, *Erkenntnis*, *Teorema*, *Dialectica* y *Transactions of the Charles S. Peirce Society*. Áreas de interés: teoría del conocimiento, filosofía

de la percepción, filosofía del lenguaje, filosofía contemporánea (principalmente, pragmatismo clásico y contemporáneo) y ética.

Carolina Mahler (carolinamahler@gmail.com)

Doctoranda en Filosofía. Traductora pública nacional de inglés. Docente de Profesionales Téc., Sup., EGB3, Pol. y TTP. Directora del equipo de investigación “Análisis Lingüístico-Conceptual del Discurso Artístico”, GRFT, acreditado por MINCyT, Pcia. de Córdoba. Investigadora integrante de “Algunas variedades de contenido mental”, varios proyectos acreditados por SECyT y CONICET, dirección Kalpokas y Aguilera. CIFFyH, Universidad Nacional de Córdoba. Profesora asociada en LEFE 1, 2 y 3 en el C.C.C. en la Lic. en Interpretación Musical, FAD, Universidad Provincial de Córdoba. Profesora titular de fonética del inglés, Colegio Universitario de Periodismo “Obispo Trejo y Sambraia”.

Mauricio Martínez (martinez_MS75@hotmail.com)

Licenciado en Ciencia de la Educación y en Psicología. Magíster en Psicología Cognitiva y Aprendizaje. Profesor en la carrera de Psicología en la UAI (asociado por concurso) y la UBA (ayudante de primera). Profesor en las maestrías de Psicología Cognitiva y Aprendizaje de la FLACSO y de Psicología Cognitiva de la UBA. Presidente de la Asociación Argentina de Profesionales del Espectro Autista (AAPEA). Director del departamento de Investigación y Docencia de APadeA. Sus trabajos de investigación se focalizan en el desarrollo perceptivo e intersubjetivo durante el primer año de vida, la interacción adulto-bebé, el desarrollo del juego en la infancia, y la intervención temprana

en autismo. Recientemente, publicó el libro *Intervención psicoeducativa para niños con TEA*, por la editorial Miño y Dávila.

Alejandra Beatriz Menti (alebmenti@gmail.com)

Doctora en Ciencias del Lenguaje con mención en Lingüística Aplicada, en la Universidad Nacional de Córdoba, Argentina. Profesora en la cátedra Lectura y escritura en lengua materna y lengua extranjera en la Facultad de Lenguas (UNC). Miembro del cuerpo docente en la Maestría en Español como Lengua Extranjera y en la Maestría en Inglés de la Facultad de Lenguas (UNC). Investigadora del Consejo Nacional de Investigaciones Científicas y Técnicas de Argentina (CONICET). Su área de interés es el desarrollo lingüístico y cognitivo infantil. Específicamente, analiza las oportunidades de aprendizaje de vocabulario que brindan las situaciones de enseñanza en clases de ciencias sociales y naturales de jardín de infantes y de la escuela primaria. Su estudio se focaliza, especialmente, en el análisis de las secuencias interaccionales entre la maestra y los niños que incluyen narrativas, en cuanto estas secuencias podrían constituir un anclaje para la elaboración de conceptos sobre el entorno natural y social y el aprendizaje de vocabulario, el establecimiento de relaciones semánticas y la realización de inferencias sobre el significado de las palabras. Ha publicado, en coautoría, libros de actividades para el aula, materiales pedagógicos, capítulos de libros y artículos en revistas especializadas.

Fernando G. Rodríguez (fgrxyz@gmail.com)

Psicólogo, magíster en Psicología Cognitiva y doctor por la Facultad de Filosofía y Letras de la Universidad de Buenos Aires. Es profesor titular de Semiología y comunicación

humana y profesor asociado de Filosofía de la Universidad Abierta Interamericana (UAI), y, asimismo, profesor asociado del Depto. de Ciencias Sociales y Humanidades de la Universidad Argentina de la Empresa (UADE). Se especializa en teoría del signo y en el desarrollo de habilidades semióticas en la infancia temprana. Ha publicado, en calidad de editor, *El estructuralismo en sus márgenes. Ensayos sobre críticos y disidentes* (Del signo, 2011) y, en calidad de traductor y anotador, *Sigmund Freud. Textos inéditos y documentos recobrados* (2017-2018), ambos en colaboración con Mauro Vallejo, así como artículos en revistas nacionales y extranjeras. Ha dirigido investigaciones teóricamente centradas en la interfaz intersubjetividad-semiosis-desarrollo ontogenético.

Celia Renata Rosemberg (crossoem@hotmail.com)

Doctora de la Facultad de Filosofía y Letras de la Universidad de Buenos Aires. Profesora titular de la cátedra de Investigación Educacional de la Facultad de Filosofía y Letras de la Universidad de Buenos Aires. Investigadora Principal del CONICET (Consejo Nacional de Investigaciones Científicas y Técnicas de Argentina). Sus proyectos de investigación abordan el desarrollo lingüístico y cognitivo infantil de niños de distintos grupos sociales y culturales. Actualmente, se desempeña como directora del CIIPME (Centro Interdisciplinario de Investigaciones en Psicología Matemática y Experimental). Ha publicado, en coautoría, libros de actividades para el aula, materiales pedagógicos, capítulos de libros y artículos en revistas especializadas.

Zoé Sánchez Barbieri (zoesanchezb@gmail.com)

Es licenciada en Psicología y doctoranda en Filosofía por la Universidad Nacional de Córdoba. Sus principales áreas de interés y estudio son la filosofía de la mente, la psicología cognitiva y la neurociencia. Actualmente, posee una beca doctoral otorgada por CONICET en la que investiga las habilidades que poseen los individuos para reconocer y atribuir emociones. También, participa del proyecto de SECyT “Intencionalidad en animales humanos y no humanos: atribución y contenido”, y ha realizado numerosos trabajos acerca de las habilidades de atribución psicológica, la naturaleza y el contenido de las emociones.

Carolina Scotto (carolinascotto@gmail.com)

Doctora en Filosofía. Profesora titular regular de Filosofía del Lenguaje de la UNC e investigadora independiente de CONICET. Especialista en temas de filosofía de la mente y del lenguaje, tales como atribución psicológica y cognición social, en humanos y animales, en enfoques interactivos de la atribución psicológica y semántica y en la filosofía del segundo Wittgenstein. Ha dirigido proyectos de investigación, tesistas de doctorado, carreras de posgrado y ha desempeñado numerosas funciones académicas y de gestión en la UNC.

